12-10-2014

Environmental Genomics and Transcriptomics of Salpa thompsoni and Population Genetic Variation of Euphausia superba in the Southern Ocean

Paola G. Batta Lona
University of Connecticut - Avery Point, paola.batta_lona@uconn.edu

Environmental Genomics And Transcriptomics Of *Salpa thompsoni* And Population Genetic Variation Of *Euphausia superba* In The Southern Ocean

Paola G Batta-Lona, PhD

University of Connecticut, 2014

A critical challenge for scientist and environmental managers is to understand how plankton populations will respond to the increase in temperature in the Southern Ocean. The Western Antarctic Peninsula (WAP) region, which is experiencing increased sea surface temperatures, is a highly productive region of the Southern Ocean, supporting dense populations of the salp *Salpa thompsoni* and krill *Euphuasia superba.* The region has complex bathymetrically-driven current flow dominated by the eastward flow of the Antarctic Circumpolar Current (ACC), which intrudes onto the shelf through deep channels, and gives rise to heterogeneous biological and chemical environments that vary temporally and spatially. *S. thompsoni* and *E. superba* are important grazers in the pelagic Southern Ocean plankton community. Ecological data suggest that important species like salps and krill are being affected by climate change and that this effects on these planktonic species will have considerable consequences for the balance of the Southern Ocean ecosystem.

The Antarctic krill *E. superba* is one of the best-studied marine zooplankton species in terms of population genetic diversity and structure; with few exceptions, previous studies have shown this species to be genetically homogenous at larger spatial scales. However few studies have discriminated life stages and examined sub-regional scale population genetic variation.

The Southern Ocean *S. thompsoni* is subject to variable environmental (temperature) and biological conditions (food availability, timing of reproduction), as well as the marked seasonal variability and long-term climate change. Knowledge of the physiology and molecular biology of Antarctic marine species such as these two key species is particularly pertinent in terms of both their adaptation to living in an extreme environment and also their responses to climate change. One of the approaches that can be used is whole genome and transcriptome sequencing. The first goal of this project was to construct a draft genome and transcriptome for *S. thompsoni.* To further enable description of functional components with genetic consequences, a preliminary genome sequence was assembled. The draft genome provided the first basis for the construction of an annotated gene catalog for *S. thompsoni*, a much-needed tool to understand the physiological response to environmental condition at a molecular level. Another goal was to assemble a reference transcriptome and examine spatial patterns of gene expression in relation to environmental conditions (i.e., temperature, salinity, nutrients and chlorophyll a), with special attention to analysis of metabolic pathways and genes thought to be involved in environmental stress response. Both genome and transcriptome annotation for *S. thompsoni* matched sequences described for marine model species such as *Ciona intestinalis* (Tunicata), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata). All 2-way ANOVAs showed clear differentiation of gene expression patterns for a wide diversity of genes in salp samples from different regions, seasons, and samples. Stress-response genes were identified and showed different patterns of gene expression between seasons.

Lastly, sub-regional scale population genetic structure of *E. superba* using SNP as molecular markers examined hypotheses of the source(s) of recruitment for krill population in the Western Antarctic Peninsula (WAP). Analysis of molecular variation revealed no significant

differentiation between 2001 and 2002 samples considered together, although 2001 and 2002 collections from Marguerite Bay showed small but significant differentiation. Larger levels of differentiation were observed among samples when furcilia larval stages were analyzed separately, and among life stages in the 2001 collection from Marguerite Bay. Explanation for the observed patterns of population genetic diversity and structure of *E. superba* populations in the WAP region is multiple centers of reproduction and recruitment, utilizing coastal, shelf, and offshore areas.

Environmental Genomics And Transcriptomics Of *Salpa thompsoni* And Population Genetic Variation Of *Euphausia superba* In The Southern Ocean

Paola G Batta-Lona

B.S., Universidad Autonoma de Baja Califronia, Mexico, 1999

M.S., University of Connecticut 2005

A Dissertation

Submitted in Partial Fulfillment of the

Requirement for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2014

APPROVAL PAGE

Doctor of Philosophy Dissertation

Environmental Genomics And Transcriptomics Of *Salpa thompsoni* And Population Genetic Variation Of *Euphausia superba* In The Southern Ocean

Presented by

Paola G. Batta-Lona, B.S., M.S.

Major Advisor_____
Dr. Ann Bucklin

Associate Advisor_____
Dr. Hans G. Dam

Associate Advisor_____
Dr. Kent Holsinger

Associate Advisor_____
Dr. Rachel O'Neill

Associate Advisor_____
Dr. Sigrid Schiel

Associate Advisor_____
Dr. Peter Wiebe

University of Connecticut
2014

Acknowledgements

I want to thank all of the people and colleagues who helped me during my doctoral studies, first my advisors and committee members.

Ann Bucklin, thank you very much for all the networking and field opportunities, every meeting, international class and research expedition. You have opened possibilities I never imagined to be within reach. They have enriched my personal and professional life. Thank you very much for believing in me and for helping me to complete both my masters and doctoral degree.

Peter Wiebe for sharing all your Antarctic knowledge and experience with me on and off the boat. I owe all of my MATLAB skills to you; your patience to teach and help me debug my mfiles helped me to get through. I would also like to thank the rest of the associate advisors, Dr. Hans Dam, Dr. Kent Holsinger, Dr. Rachel O'Neill and Dr. Sigrid Schiel, all of your valuable scientific insight and guidance helped me shape the content and significance of my doctoral research.

Many of the analyses for my doctoral research required different sorts of expertise; I would like to thank Craig O., Nate J., Claudia K., Leocadio B. and Amy M. for sharing their expertise from cruise planning to analyzing samples.

I would like to thank my family, Lupita, Miguel, Joao, los chuchines, mi abuelita Luz y Faride who have been supportive since I left home from my undergraduate studies. To my CT-Mexican family, Xochitl and Marek, who adopted me and cared for me all these years. And last but not least to all the Repsher family in particular to Will and Ginger for your constant

willingness to help. Your unconditional love, presence and support got me through stressful and joyful times.

Thank you to the Marine Sciences Department of the University of Connecticut, including the administrative staff, diving and shop staff. Everybody was always ready to help and make my research needs available. I particularly want to thank the continued support and motivational words of: Mrs. Pat Evans, Mrs. Deb Schuler and Mrs. Elise Hayes.

Finally I would like to thank the National Science Foundation for supporting my research (Award No. ANT-1044982). The Marine Sciences Department of the University of Connecticut has also provided financial support through internal grants.

Table of Contents

Introduction

Background:

A critical challenge for scientists and environmental managers is to understand how plankton populations will respond to the increase in temperature in the Southern Ocean. The Western Antarctic Peninsula (WAP) region is known to support large standing stocks of the Southern Ocean salp *Salpa thompsoni* and the Antarctic krill *Euphausia superba* (Marr, 1962; Siegel, 2005; Atkinson et al., 2008). The WAP region, one of the most biologically-rich areas of the Southern Ocean, is one of the most rapidly warming regions on Earth (Clarke et al., 2007; Ducklow et al., 2007; Loeb et al., 2008). Winter air temperatures are increasing and the duration of winter sea ice cover is decreasing, due to later advance and earlier retreat of the ice (Ducklow et al., 2007; Yin et al., 2011), although winter ice coverage around the Antarctic continent has reached record amounts in the past few years (Reid et al., in press). However, if there are changes in heat flux, this will affect the mixed layer in the Antarctic Circumpolar Current (ACC) and as a consequence impact the vertical flux of nutrients and limiting elements (Flores et al., 2012).

Both species are subject to severe environmental (temperature) and biological conditions (food availability, energetic constraints, timing of reproduction), as well as the marked seasonal variability and are vulnerable to climate change of the Southern Ocean ecosystem.

Salpa thompsoni is typically found in warmer, less productive water than krill, and the species tends to be more abundant following winters when the sea ice cover is less extensive (Loeb et al., 1997). The extent of sea ice appears to be important in determining community structure – either because the algae it contains may be an important food source for the overwintering krill or because it may protect the krill from potential predators (Siegel, 2005).

Euphausia superba is a swarming crustacean, whose discontinuous distribution and highly variable time/space patterns of concentration have long been a subject of research (Marr, 1962; Spiridinov et al.,1996; Lawson et al., 2008). Both reproducing and non-reproducing krill are consistently found in the WAP region and krill may be retained in this region by persistent gyres situated over the northern shelf (Marr, 1962; Lascara et al., 1999). For *E. superba*, Zane et al. (1998) demonstrated significant temporal variation of populations based on DNA sequence analysis of the mitochondrial NADH dehydrogenase (ND-1) gene. SNP identification, screening, and analysis have become increasingly rapid and inexpensive, with the use of relatively inexpensive SNP detection kits and standard protocols (Quintans et al., 2004).

Salpa thompsoni (along with krill) is one of the major grazers in the Southern Ocean. Salps are highly efficient, non-selective filter-feeders, capable of ingesting particles between 1-1000 μm (Madin and Deibel, 1998). The repackaging of waste material in large, fast sinking fecal pellets means that *S. thompsoni* is a major source for vertical flux of organic matter (Huntley, 1989; Dubischar and Bathmann, 1997; Le Fèvre et al., 1998; Perissinotto and Pakhomov, 1998a, 1998b; Pakhomov et al., 2002; Pakhomov, 2004; Pakhomov and Froneman, 2004; Phillips et al., 2009).

Along the WAP region, the phytoplankton community is already thought to have changed due to impacts of climate change (Montes-Hugo et al., 2009). The phytoplankton community in the WAP region appears to be decreasing and shifting from larger diatoms (15-270 μm) to smaller cryptophytes (8 ± 2 μm) (Moline et al., 2004). The shift in phytoplankton biomass and size has direct consequences for grazer communities, especially Antarctic krill (*E. superba*), which are inefficient at grazing small cells, whereas tunicates such as the salp (*S. thompsoni*) are efficient at grazing the smaller cells (Moline et al., 2004).

Salpa thompsoni frequently forms dense blooms covering large areas of the Southern Ocean during the austral summer. Salps alternate between morphologically different asexual and sexual forms. The asexual form (oozooid) is shaped like a cylinder and has a stolon that forms chains (clonal aggregate salps) that are eventually released into the water column when the conditions are favorable (i.e., austral spring and summer) (Loeb. et al., 2012). The stage of the stolon development is used to characterize the life stage of the oozooids. Studies characterizing the different life stages of *S. thompsoni* have aided understanding of population dynamics and life cycle adjustments in response to environmental variability (Chiba et al., 1999). This life cycle strategy facilitates rapid population increase when environmental conditions are advantageous and can often result in dense salp blooms (Alldredge and Madin, 1982). During "salp years", *S. thompsoni* essentially replaces Antarctic krill in the waters off the Antarctic Peninsula (Huntley et al., 1989; Park and Wormuth, 1993; Nishikawa et al., 1995; Siegel, 1995; Loeb et al., 1997). The oceanography and zooplankton community composition have been documented for both the WAP region and the Indian sector of the Southern Ocean. However few studies have shown comparisons between temporal and spatial patterns of environmental conditions with respect to salp presence and absence. Regardless of the mechanism, warming of the Southern Ocean appears to have favored an increase in *S. thompsoni* distribution and abundance over Antarctic krill, *E. superba* (Atkinson et al., 2004).

Knowledge of the physiology and molecular biology of Antarctic marine species is essential to understand their adaptation to living in an extreme environment and also their responses to climate change. High throughput sequencing has revolutionized the study of whole-genome (i.e., study of the genomes of organisms, including intensive efforts to determine the entire DNA sequence of organisms and fine-scale genetic mapping efforts), and whole-transcriptome (i.e.,

patterns of gene expression based on RNA presence and quantity from a genome at any given time). The advances in next generation sequencing have made whole-genome and whole-transcriptome sequencing more accessible and available for non-model organisms. *Ciona intestinalis* (Dehal et al., 2002) and *Oikopleura dioica* (Seo et al., 2001) are the closest relatives to *S. thompsoni* for which whole-genome sequencing has been completed and for which the information is available. On the other hand, whole-transcriptome (RNA-Seq) studies are also providing insights into how organisms are coping with different environmental conditions (Gallardo et al., 2014; Villarino et al., 2014; Smith et al., 2013). Therefore these tools can be used to answer ecological questions and to understand the effects of different environmental conditions on the zooplankton communities experiencing seasonal or long-term variability in the environment.

Research objectives:

The general objective of this dissertation is to investigate and understand distribution of *S. thompsoni* and *E. superba* in the Southern Ocean, with an emphasis on molecular responses of *S. thompsoni* to different environmental conditions. This research has provided genomic and transcriptomic data for a previously uncharacterized representative of the Salpidae (an important tunicate group) and will also provide quantification of sub-regional scale population genetic diversity and structure of the Southern Ocean krill, *E. superba,* in the WAP region.

This dissertation is organized into four chapters addressing the following objectives:

Chapter 1. Reference genome of the Southern Ocean salp, *Salpa thompsoni*, in the Western Antarctic Peninsula

The goal of this chapter is to create new genetic and genomic tools for analysis of the physiological processes associated with life history events and response to environmental conditions of *Salpa thompsoni* in the Southern Ocean.

Chapter 2. Multivariate analysis of environmental conditions associated with presence of the Southern Ocean salp, *Salpa thompsoni,* in the Western Antarctic Peninsula region.

The goal of this chapter is to describe similarities and differences in *S. thompsoni* habitats between regions and seasons with different environmental conditions, and provide new insight into environmental conditions associated with salp distribution in the sampled Southern Ocean regions.

Chapter 3. Transcriptome-wide profiles of gene expression of *Salpa thompsoni* in relation to variation of the pelagic environment of the Southern Ocean

The goal of this chapter is to provide new insights into the relationship between environmental variability and patterns of gene expression of *S. thompsoni* and to begin to identify important genes and gene families associated with adaptation to the unique environmental conditions of the Southern Ocean.

Chapter 4. Population genetic variation of the Southern Ocean krill, *Euphausia superba*, in the Western Antarctic Peninsula region based on mitochondrial single nucleotide polymorphisms (SNPs)

The goals of this chapter are to examine sub-regional scale population genetic diversity and structure of *E. superba* using molecular characters selected with this goal in mind, and to thereby examine hypotheses of the source(s) of recruitment for krill populations of the Western Antarctic Peninsula region.

Chapter 1:

Reference Genome of the Southern Ocean salp, *Salpa thompsoni*

ABSTRACT

A preliminary genome sequence has been assembled for the Southern Ocean salp, *Salpa thompsoni* (Urochordata, Thaliacea). The draft genome for this salp species is one of the few available genome sequences for urochordates, an ecologically important and evolutionarily divergent group of organisms. Up to now, only two of the families of urochordates have had their genome sequenced: *Ciona intestinalis* and *Oikopleura dioica*. Despite being recognized as a key component of Antarctic pelagic ecosystems, no study has been published on *S. thompsoni* population genetics or genomics, reflecting the lack of genetic studies conducted on salps and other thaliaceans. Although the work described here is the first of its kind on a salp species, the molecular genetic techniques are now being widely used on many other organisms to answer a wide variety of questions. The new data presented here provides the basis for the construction of the first annotated gene catalog for *S. thompsoni*. Moreover, this study provides a much-needed tool to enable description of functional components within the *S. thompsoni* genome and to understand their physiological responses to environmental conditions at a molecular level.

1. INTRODUCTION

The Phylum Chordata includes three sub-phyla: Vertebrata, Cephalochordata, and Tunicata. The tunicates include three classes of organisms: Ascidiacea, Thaliacea, and Larvacea. The most studied tunicates are the ascidians, including the model species, *Ciona intestinalis* (Holland et al., 2003, Takatori et al., 2004; Kawada et al., 2011). Larvaceans have also been the subject of genomic analysis, including *Oikopleura dioica* (Seo et al., 2001; Stach, 2007; Yadetie et al., 2012). The species representatives for these two groups have been extensively studied for

evolutionary and developmental biology, in part because they are easily cultured in the laboratory. A number of studies have used them to study the development of chordate innovations, such as the tail, notochord, and dorsal nervous system.

The third group of urochordates, the thaliaceans, comprises a group of pelagic zooplankton that is widely distributed across the world oceans. Some species can occur in dense populations during bloom conditions in high latitudes (Daponte et al., 2001; Pakhomov et al., 2011) or in water with high productivity (Deibel et al., 2009; Wiebe et al., 1979). However members of this group have not been successfully cultured in the laboratory. As a result, thalaceans are not as well studied as the other two groups.

Salps (*Salpa thompsoni*) belong to the order Salpida, one of three groups within the class Thaliacea. Their life history alternates between asexual and sexual stages. A sexually-produced solitary individual, termed oozoid,forms chains (blocks of identical, clonal aggregate salps) that develop from stolons through asexual reproduction, which are eventually released. A solitary salp may produce 4-5 blocks totaling over 1,500 individual aggregates. *Salpa thompsoni* is a protogynous hermaphrodite; all of the members of the newly released chain (aggregate) are females. The members of an asexually-produced chain (aggregate females) are fertilized and then nurture developing solitary embryos. Each aggregate produces a single embryo, once this embryo is released and the life cycle has been completed, the female aggregate will turn into a male. This life cycle facilitates rapid population increase when environmental conditions are advantageous, and can often result in dense salp blooms (Alldredge and Madin, 1982).

High throughput sequencing has revolutionized the study of genomics (i.e., study of the genomes of organisms, including intensive efforts to determine the entire DNA sequence of organisms and fine-scale genetic mapping efforts), including the analysis of non-model (Vera et al., 2008; Johansson, 2009) and marine organisms (e.g., Meyer et al., 2009; Quin et al., 2008). The work described herein provides new genetic and genomic tools for analysis of the physiological processes associated with life history events and adaptation to environmental conditions of *S. thompsoni* in the Southern Ocean.

2. METHODS

2.1 Sample Collection

Collections of *S. thompsoni* were made in January, 2009 on the Japanese research vessel Umitaka-Maru in the Indian sector of the Southern Ocean. The cruise transect was from Cape Town, South Africa to Fremantle, Australia. Samples were collected at Station 33 (Fig. 1) at depths of 50 -100 m with a Rectangular Midwater Trawl. Once the salps arrived on board, the samples were analyzed under the microscope to confirm species identification, and standardized measurements (such as length and life stage determination) were performed. The gut of *S. thompsoni* was removed by dissection to avoid DNA contamination from the prey, and muscular tissue was flash frozen with liquid nitrogen and preserved at -80^{o} C. Four salps (life stage 2) were chosen, their DNA was extracted following the QIAGEN (Germantown, MD) DNeasy extraction protocol (Cat. no. 69582) and RNA purified prior to sequencing the samples.

2.2 Genome Size Estimation

Four salps were selected for analysis and their DNA was extracted in the same matter as described in the previous section. We used real-time quantitative PCR (Wilhelm et al. (2003) to estimate the genome size for *S. thompsoni*. A single copy gene, *Tbx1*, found in *Ciona intestinalis* (Takatori et al., 2004) was identified within *S. thompsoni* and validated for use as a reference standard.

2.3 Next Generation Sequencing - 454 and Ion Torrent Platforms

2.3.1 454 GS FLX system

Four different *S. thompsoni* individuals collected in the Indian Ocean were used for genome sequencing. Three specimens were used for library construction and subsequent sequencing using shotgun and paired-end protocols provided by the instrument manufacturer (GS FLX Titanium General Library Preparation Guide; Roche Applied Science, Branford, CT, USA).

Shotgun DNA library generation and sequencing: Salp genomic DNA was mechanically sheared into fragments to which specific adaptors necessary to the GS FLX sequencing process were ligated. After adaptor ligation, the fragments were denatured and clonally amplified via emulsion PCR (emPCR), generating millions of copies of template per bead. The DNA beads were then distributed into picoliter-sized wells on a fiber-optic slide (PicoTiter-Plate™), along with a mixture of smaller beads coated with the enzymes required for the pyrosequencing reaction. The four DNA nucleotides were then flushed sequentially over the plate. Light signals were released upon base incorporation and captured by a CCD camera, and the sequence of bases incorporated per well was stored as a read.

Paired-end DNA library generation and sequencing: For the paired-end libraries, genomic DNA from a third individual was sheared into ~3 kb fragments, and biotinlylated hairpin adaptors

(containing an *EcoRI* site) were ligated to the fragment ends. The fragments were subjected to *EcoRI* digestion and circularized by ligation of the compatible ends, and subsequently randomly sheared. Biotinlyated linkers containing fragments were isolated by streptavidin-affinity purification. These fragments were then subjected to 454 sequencing on the GS FLX system. The paired end reads were recognizable as the known linker (originating from the two hairpin adaptors). When sequenced on the GS FLX, this protocol generated two, ~100 bp tags known to be ~3 kb apart. These paired end reads were used to aid in the assembly of contigs into scaffolds.

2.3.2 Ion Torrent Personal Genome Machine (PGM)

The fourth specimen was sequenced twice with the Ion Torrent™ PGM (Life Technologies, Grand Island, NY). The Ion Torrent PGM uses semiconductor technology to detect the protons released as nucleotides are incorporated during synthesis (Glenn, 2011). Library construction and the emPCR were performed according to the Ion Torrent protocol associated with the kit (Ion Xpress™ Fragment Library Kit).

2.4 Filtering and Assembly

Sequencing reads generated on both the Roche 454 System and the Ion Torrent PGM were used in developing a *de novo* genome assembly. Initially, sequencing reads were filtered based on phred quality scores (Q) below 30 (Q30 is defined as the probability that a given base is called incorrectly is 1 in 1000). Known bacterial or algal genome sequences were excluded from all further analyses as follows. If a sequencing read showed 90% similarity to a bacterial or algal genome over 90% of the read length, the read was tagged as a possible contaminant and removed from the dataset. Due to the innate error occurrence in sequencing analysis in the Proton

platform, all Proton reads were error-corrected using the program Coral v1.4 with settings of mr=2, mm=2, and g=3 (Salmela and Schröder 2011). The remaining data were assembled into contigs and scaffolds using the assembly program Newbler v2.8. The subsequent assembly was then further scaffolded using the program L-RNA Scaffolder (Xue et al., 2013), which scaffolds contigs using a BLAST-mapped transcriptome reference. The transcriptome sequence used in the study was generated from sequencing runs for RNA-Seq from multiple individuals (N=40) that were then normalized and assembled using the program Trinity and associated scripts (see chapter 3).

2.5 Annotation

The final genome assembly was annotated by generating *de novo* predictions for both repetitive elements and gene sequences. Repetitive elements were predicted using the programs RepeatModeler (Smit, 2008-2010) and RepeatMasker (Smit, 1996-2010) to identify both *de novo* and reference-based repetitive elements, respectively. All unclassified repetitive elements were further examined using the program TEClass (Abrusán et al., 2009) for a more detailed annotation assignment. The genome sequence was then assessed for repetitive element content using RepeatMasker and our *de novo* repeat annotations. Gene and protein sequences were predicted using the Maker2 (Holt and Yandell, 2011) pipeline, in which repetitive elements were masked and salp transcriptome sequences and *Ciona intestinalis* proteins for which the salp transcriptome showed at least 90% coverage were aligned to *S. thompsoni* genome to generate gene predictions. Subsequent iterations of the program used two *de novo* gene prediction programs, SNAP (Korf, 2004) and Augustus (Stanke and Morgenstern, 2005), to generate additional *de novo* gene predictions. Both SNAP and Augustus were trained for *S. thompsoni* using full-length protein sequences isolated from the transcriptome; these proteins were

identified by comparing their predicted open-reading frame sequences to those of existing proteins in the UNIPROT database (http://www.uniprot.org/). Ultimately, both evidence-based and *de novo* predictions were used to cross-validate and finalize gene predictions. All predicted proteins were compressed into a non-redundant protein set using the CD-Hit algorithm with default parameters (Fu, et al., 2012) and then annotated by identifying their protein identity using BLASTP with e-value of 1E-5).

2.6 Genome Statistics

Gene Ontology terms were assigned to the *S. thompsoni* genome with the BLAST2GO B2G4 v2.5 pipeline using default parameter settings (Conesa et al., 2005) and protein family and domain information using InterProScan v.5.6 (e-value = 1E-5). Enzyme codes and KEGG pathway membership were all done in BLAST2GO as well. All final annotations were further summarized in GO SLIM (http://geneontology.org/page/go-slim-and-subset-guide).

3. RESULTS

3.1 Genome Size Estimation

Averaging across four different salp samples, genome size was estimated to be 602 Mb (± 173 Mb), indicating that the depth of coverage for this sequencing effort (11.8 Gb) was ~20X, based on a total read length of approximately 12 Gb.

3.2 Next Generation Sequencing and Assembly

After post-filtering for possible contaminants, the genome assembly efforts utilized 881,748 paired end reads with 3kb inserts and 1,845,108 single end reads from 454 sequencing and 106,495,320 reads from Ion Proton sequencing (109 million reads or 11.8 Gb total). Of these

reads, 12,748,500 were discarded because they were too short, belonged to a simple repeat, or were of low quality (as determined by Newbler v2.8 software). Overall, the total amount of sequence input into the assembly was 10 Gb from 96,037,016 reads.

The initial assembly yielded 859,784 contigs accounting for 327 Mb of sequence, with a large portion of the reads utilized in the assembly. Further scaffolding of these contigs resulted in a final draft sequence build of 478,293 scaffolds/contigs with an N50 value (i.e., the minimum size of 50% of the contigs; Kasahara et al., 2007) of 934 bp and containing 318 Mb in the sequence build with a GC-content of 35.92%. As indicated by the number of contigs and low N50 values, although the sequence likely covers the majority of the genome, 43% of the reads were discarded, and therefore this build is still somewhat fragmented.

3.3 Annotation

The CEGMA (Core Eukaryotic Genes Mapping Approach) coverage of core eukaryotic genes reflects the somewhat fragmented genome build, indicated by the number of contigs and low N50 values, with only 35% of the genes found to be represented in their complete sequences, while 76% of the genes showed as least partial coverage in our genome build (Table 1). Repetitive elements accounted for 33.06% of the assembled sequence with DNA transposons, unclassified novel repetitive elements, long interspersed nuclear elements (LINEs), and long terminal repeats (LTRs) making up the majority of that percentage (17.59%, 6.55%, 4.61% and 2.82% of the total sequence, respectively). Gene prediction generated 34,681 cross-validated gene predictions and 29,752 non-overlapping *ab initio* gene predictions. Together, these predictions yielded 31,882 non-redundant gene predictions. The average protein length for these gene predictions was 161 residues; 29, 264 of these proteins were 50 residues or longer. A total

of 13,454 of the proteins could not be identified by homology to any other protein sequence in the NCBI non-redundant (NR) database (http://www.ncbi.nlm.nih.gov/refseq/). Of these, however, 3,190 proteins had identifiable functional domains. Moreover, 15,810 proteins carried significant similarity to existing proteins and could be associated with an inferred gene identity (Fig. 2). High identity matches for *S. thompsoni* genome sequence were primarily from the model species *Ciona intestinalis* (Tunicata), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata) (Fig. 3). These results indicate the transcriptiome-guided annotations of our genome sequence data are reliable and match available data for the most closely-related model species, *Ciona intestinalis*.

3.4. Gene Ontology

Gene ontology is used to ensure consistent description of gene products across genome studies. The properties of gene products cover three categories: *Cellular component*, parts of a cell (i.e., nucleus or endoplasmic reticulum); *Molecular function*, elemental activities of a gene product at the molecular level (i.e., binding or catalysis); and *Biological Process*, molecular events or operations related to the functioning of cells, tissues, organ, and organisms (i.e., signal transduction, or more specifically, pyrimidine metabolic processes) (http://geneontology.org/page/ontology-documentation). Each category of gene ontology has a different level of detail. The *S. thompsoni* genome data was widely distributed among the different levels of gene ontology (Fig. 4). The annotated genes included diverse proteins (e.g., heat shock proteins, myosin, ubiquitin, zinc finger protein) and enzymes in various metabolic pathways (e.g., dehydrogenases, hydrolases, kinases, oxidases, phosphorylases, polymerases, reductases, reverse transcriptases, transferases, and transposases). A variety of biological

processes (e.g., metabolic and developmental processes, growth, and reproduction) and molecular functions were represented among the identified genes (Fig. 3). At a higher resolution of analysis (Fig. 4), annotated genes included those associated with oocyte fate determination, oocyte differentiation, and gonad development. Genes found during the annotation and construction of the draft genome will be used for analysis of changes in the transcriptome of individuals collected under different environmental conditions as detailed in Chapter 2.

4. DISCUSSION

Comparatively few studies have used molecular approaches to understand how organisms may respond and adapt to environmental variability at appropriate time and space scales, ranging from short-term events to long-term climate change. Few individuals were used in this study to reduce detected intraspecific variation and thus maximize the depth of coverage achieved for the number of sequences determined. The DNA sequences obtained from our genome sequencing efforts gives a quick view of the gene profile of *S. thompsoni* during the austral summer of 2009. Almost ten years ago, genome sequencing of the tunicate *Ciona intestinalis* made a substantial contribution to the fields of evolutionary and developmental biology (Holland et al., 2003). Subsequently, the sequencing of another tunicate, *Oikopleura dioica,* shattered the idea of conservation of global architecture across genomes from different organisms in the animal kingdom (Denoeud et al., 2010). The *S. thomopsoni* genome results are the first ones for this species and the Thaliaceans. Together with the other tunicate genome data available, these new data may be expected to have major impact on our understanding of chordate evolution and developmental biology.

Genome annotation for *S. thompsoni* matched sequences described for marine model species such as *Ciona intestinalis* and *Oikopleura dioica* (Tunicata), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata). Although both tunicate model species were represented in *S. thomposoni* annotation, the *Oikopleura* genome was not among the 10 top-hot species; this is likely due to the fact that *Oikopleura* has the smallest genome known for a chordate (51 – 65 Mb), (Seo, 2001) while *Ciona* has a larger genome (160 Mb) and has deeper coverage. The database included more genes from *Ciona* to map to and therefore this species was in the first place of the top-hit species list. Based on the genome size estimation, *S. thompsoni* genome size was determined to be approximately 605 Mb. This is the first estimate for this species and indicates that salps have larger genomes than the rest of the tunicates, although the implications of this for their genome architecture are yet to be seen.

This study seeks to provide a catalog to identify genes that can be used for a wide variety of research. The reference genome provides a wealth of comparative data with which to investigate the role of developmental constraint in genome evolution. The data can be used to study how organisms respond and adapt to environmental variability at the molecular level, ranging from short-term events to long-term climate change. Additional genomic analysis should focus on genes with particular biological functions that are relevant to adaptation of *S. thompsoni* to changes in the Antarctic environment, such as temperature and food availability. Transcriptome sequencing (RNA-Seq) of salps will yield new understanding of the effects the environment has on expression of genes involved in biological function, including those associated with responses of *S. thompsoni* to environmental variability and change in the Southern Ocean.

Figure Legends:

Figure 1. (A) Map of cruise transect in the Indian Sector of the Southern Ocean (austral summer 2009) and cruise tracks. (B) Magnified section of the sampling area and station where the samples were collected (red square).

Figure 2. Pie chart showing *S. thompsoni g*enome data distribution analyzed by Blast2GO

Figure 3. Histogram with the top hit species distribution, or the species whose sequences were highly similar to *S. thompsoni*

Figure 4. GO-Level distribution of *S. thompsoni* genome data for three categories: Molecular processes (green), Molecular function (blue), and Cellular component (yellow)

Figure 5. GO distribution of the top 20 GO terms in level 3 for each category; Molecular processes (green), Molecular function (blue), and Cellular component (yellow)

Figure 6. *Salpa thompsoni* genome ontology terms for Level 3 of Biological Process

Figure 7. *Salpa thompsoni* genome ontology terms for Level 13 of Biological Process

Table 1. *S.thompsoni* genome's coverage of core eukoaryotic genes.

	Complete	Partial
Ciona int.	87.90%	94.35%
Ciona sav.	83.06%	95.56%
Oikopleura	85.48%	91.53%
Botryllus		
Salpa thompsonii	35.48%	76.21%

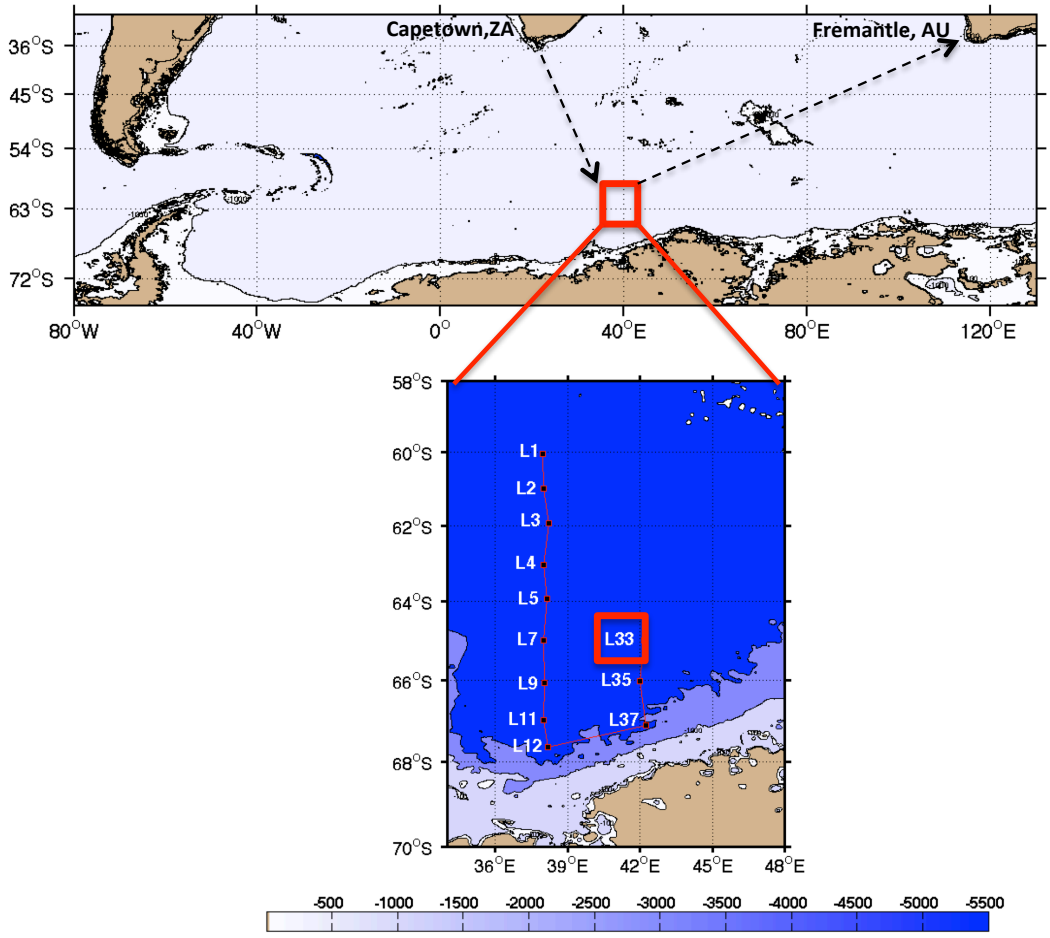


Figure 1. Map of sampling area

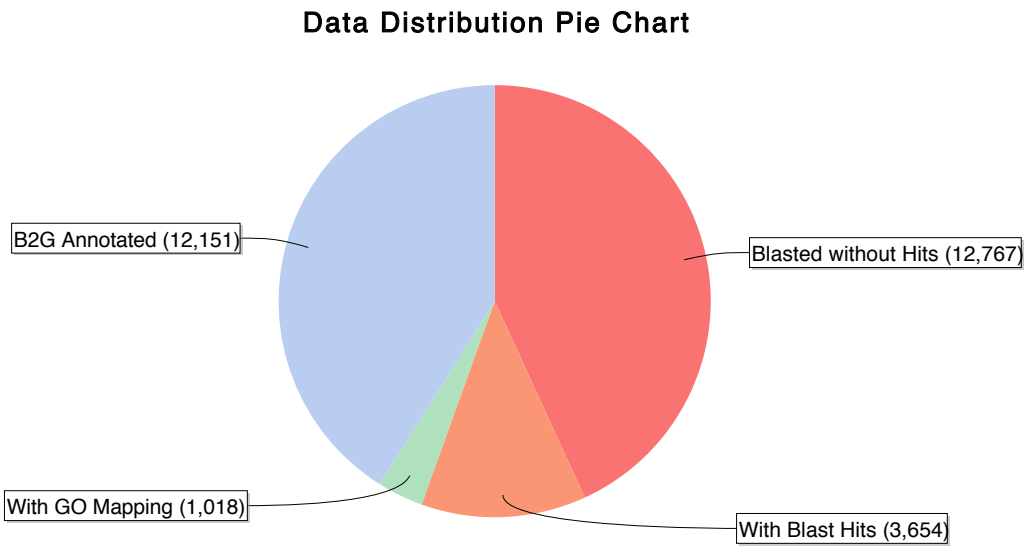


Figure 2. *S.thompsoni* genome data distribution in Blast2GO

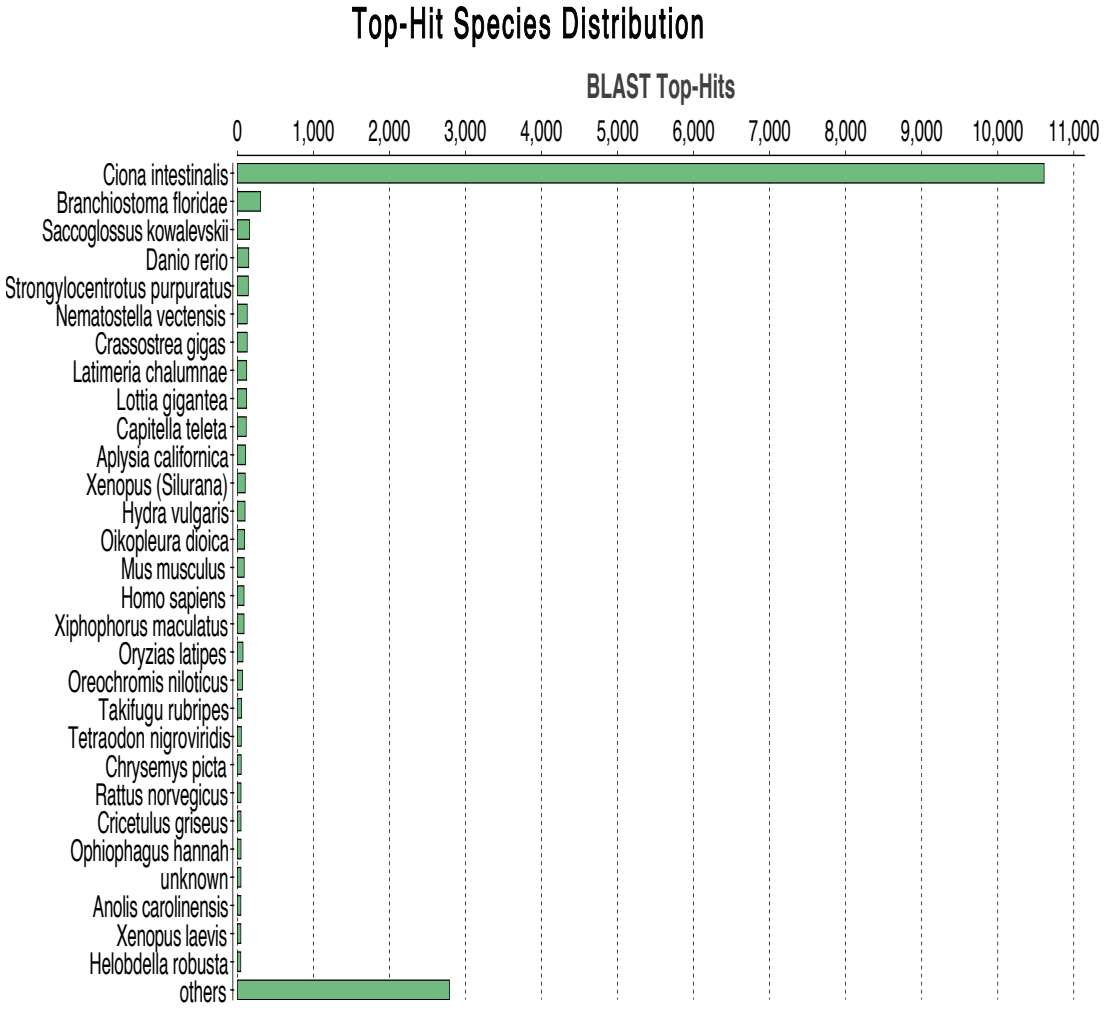


Figure 3. Top-Hit species data distribution in Blast2GO

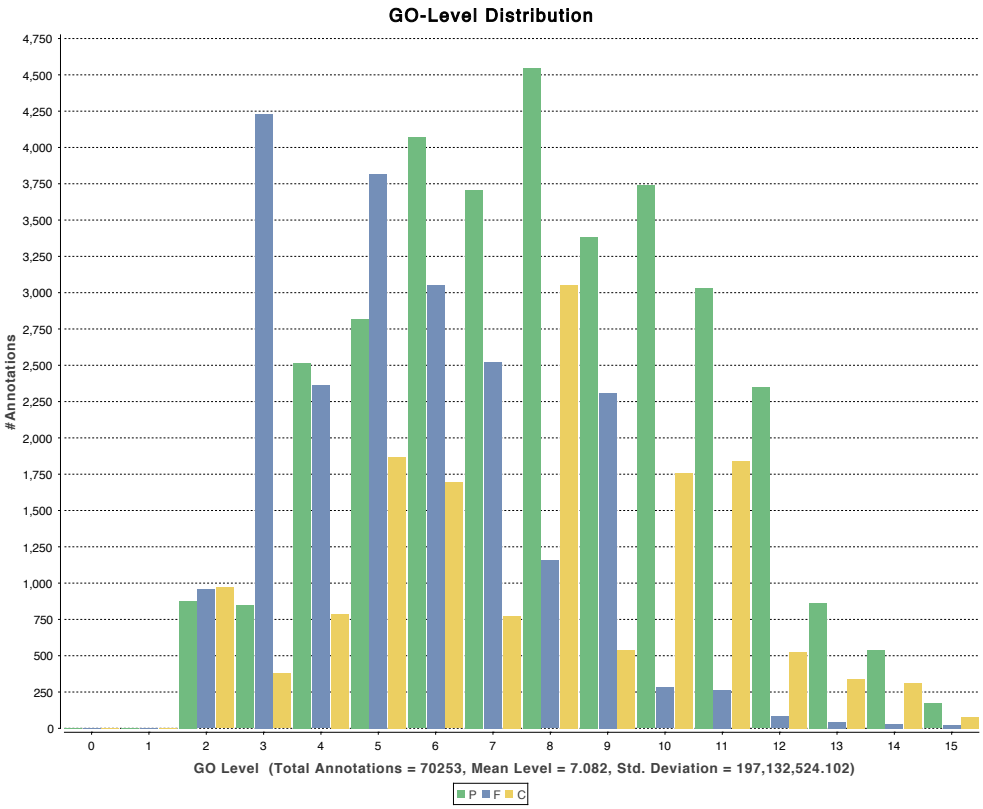


Figure 4. GO-Level distribution of *S. thompsoni* genome data

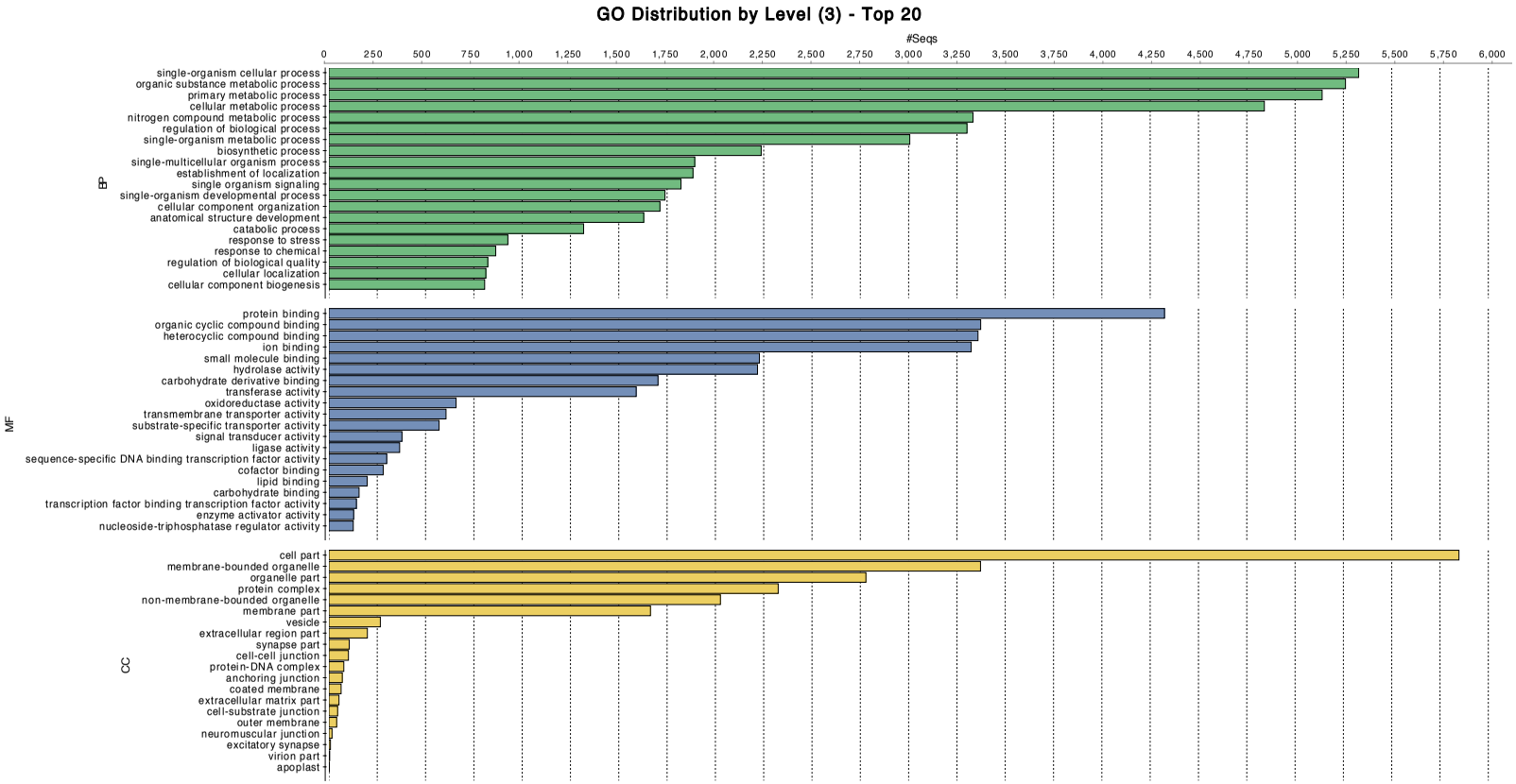


Figure 5. GO distribution of the top 20 GO terms for the three categories.

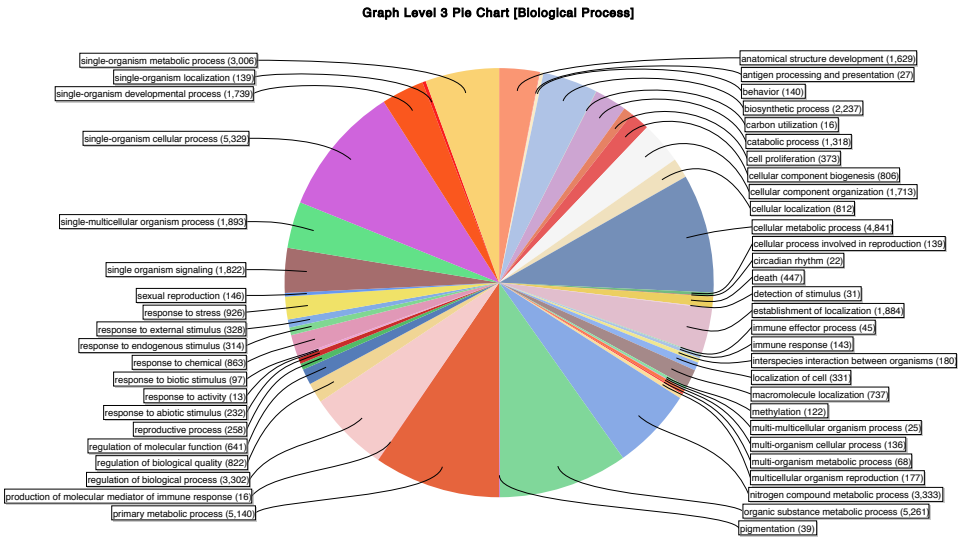


Figure 6. Ontology terms for level 3 of Biological Process

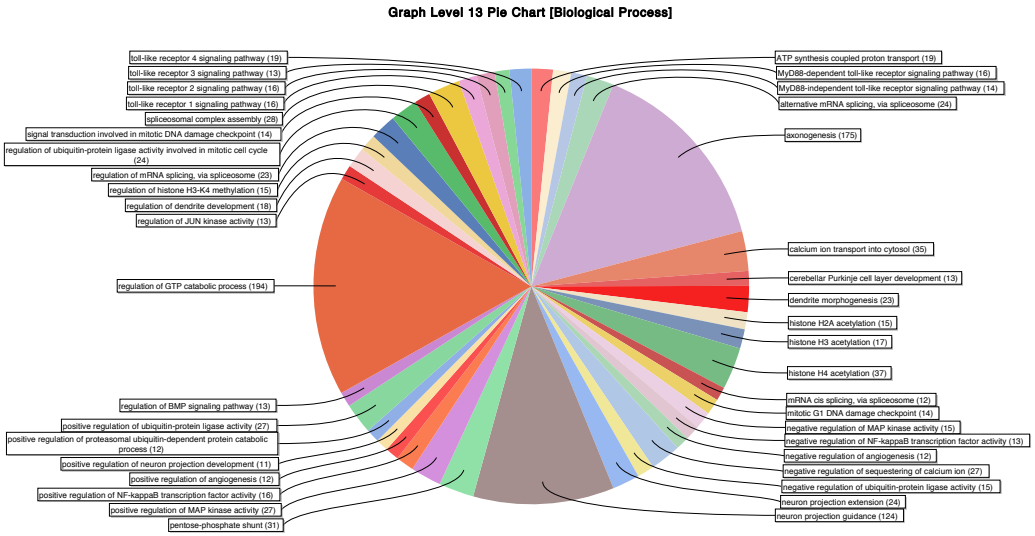


Figure 7. Ontology terms for level 13 of Biological Process

Chapter 2:

Multivariate analysis of environmental conditions associated with presence of the Southern Ocean salp, *Salpa thompsoni,* in the Western Antarctic Peninsula region and Indian Sector

ABSTRACT

The Southern Ocean salp, *Salpa thompsoni*, occurs throughout the Antarctic pelagic ecosystem, including the Western Antarctic Peninsula (WAP) region and the Indian Sector. In the WAP, a complex bathymetrically-driven current flow is dominated by eastward flow of the Antarctic Circumpolar Current (ACC), which intrudes onto the shelf through deep channels;, while the Indian Sector region is bathymetrically less complex. *Salpa thomsoni* is subject to the resulting time/space heterogeneity in physical (temperature, salinity) and biological conditions (nutrients, chlorophyll, food availability). The environmental data presented here showed the seasonal differences between spring and summer in the WAP region, including the characteristic seasonal modification of these water masses. The species shows marked variation in distribution, abundance, life stages, timing of reproduction, etc at seasonal and longer-term time scales. *Salpa thomsoni* was collected during research cruises to the WAP region during January and November, 2011, and to the Indian Sector during February, 2009. Nutrient distribution was similar in all stations, regions and seasons. PCA analysis of environmental parameters during austral spring and summer 2011 in the WAP and austral summer 2009 in the Indian Sector showed that the main factors influencing *S. thompsoni* distribution were silicate (PC1) and temperature (PC2). However when chlorophyll a was included in the PCA analysis for austral spring 2011 in the WAP and austral summer 2009 in the Indian Sector, the main factors influencing distribution of this species were chlorophyll a and temperature (PC1) and silicate (PC2). PCA analysis included in this work contributes to an understanding of the distribution of *S. thompsoni* in relation to different environmental parameters (including nutrients).

1. INTRODUCTION

The Southern Ocean surrounds Antarctica and extends northwards to 60ºS; there are multiple sub-divisions of this ocean into seas, bays, troughs, as well as two important gyres: the Ross Gyre in the Ross Sea and the Weddell Gyre in the Weddell Sea (Orsi et al., 1995). Research has been focused mostly on the western part of the Southern Ocean, while the eastern part has been less-frequently sampled (Nicol and Raymond 2012). For this study, the focus is on two areas: the Western Antarctic Peninsula (WAP) region (including offshore and shelf areas, as well as Bransfield Strait) and an area of the Southern Ocean Indian Sector south of the Enderby Abyssal plain and north of Lützow-Holm Bay (34º-46ºE and 59º-70ºS).

The WAP region, one of the most biologically-rich areas of the Southern Ocean, is also one of the most rapidly warming regions on Earth (Clarke et al., 2007; Ducklow et al., 2007; Loeb et al., 2008). Winter air temperatures are increasing and the duration of winter sea ice cover is decreasing, due to later advance and earlier retreat of the ice (Ducklow et al., 2007; Yin et al., 2011), although winter ice coverage around the Antarctic continent has reached record amounts in the past few years (Reid et al., in press). However if there are changes in heat flux this will affect the mixed layer in the Antarctic Circumpolar Current (ACC) and as a consequence impact the vertical flux of nutrients and limiting elements (Flores et al., 2012). Recently, the phytoplankton community of the WAP region appears to be decreasing and shifting from larger diatoms (15-270 μm) to smaller cryptophytes (8 ± 2 μm) (Moline et al., 2004). The shift in phytoplankton biomass and size has direct consequences for grazer communities, especially Antarctic krill (*Euphausia superba*), which are inefficient at grazing small cells, whereas tunicates such as the salp (*Salpa thompsoni*) are efficient at grazing the smaller cells (Moline et al., 2004).

When abundant, *S. thompsoni* is a dominant grazer on particulate matter and a major source for vertical flux of organic matter in the Southern Ocean (Huntley et al., 1989; Dubischar and Bathmann 1997; Perissinotto and Pakhomov 1998; Pakhomov 2004; Pakhomov and Froneman 2004). During "salp years", *S. thompsoni* essentially replaces Antarctic krill in the waters off the Antarctic Peninsula (Huntley et al., 1989; Park and Wormuth 1993; Nishikawa et al., 1995; Siegel 1995; Loeb et al., 1997). *Salpa thompsoni* is typically found in warmer, less productive water than krill, and it tends to be more abundant following winters when the sea ice cover is less extensive (Loeb et al., 1997). The extent and duration of sea ice appears to be important in determining community structure – either because the algae it contains may be an important food source for the overwintering krill or because it may protect the krill from potential predators (Siegel 2005). Regardless of the mechanism, warming of the Southern Ocean appears to have favored an increase in *S. thompsoni* distribution and abundance over Antarctic krill, *E. superba* (Atkinson et al., 2004). With the current knowledge any changes in krill abundance are considered to have a strong effect on a wide range of predators in the WAP region. However recent studies have brought forward the idea of whther salps are consumed by predators such as seabirds (Loeb et al., 2012) due to the increase of salps in higher latitudes.

The goal of the work presented here is to contribute to the genome research effort in the tunicates providing new molecular tools to the understanding of the underlying factors dictating or changing *S. thompsoni's* geographical distribution. In particular, this study seeks to characterize the biological and physical conditions of several Southern Ocean regions, including the Western Antarctic Peninsula region and Indian Sector, where zooplankton samples containing *S. thompsoni* were collected. Temporal and spatial patterns of environmental variation in the sampled regions are characterized and analyzed in comparison with salp presence/absence using

multivariate statistics, including principal component analysis. The resulting new information will further identify and characterize the environmental conditions under which salps are present in Southern Ocean waters, and thereby work toward improved understanding and eventually prediction of salp population dynamics, including bloom formation.

1.1 Description of Study Area

The Western Antarctic Peninsula (WAP) region extends north from 75º S, 80º W to the northern tip of the peninsula near 63º S, 60º W (Fig 1-A and B). Based on bathymetry, oceanography, and biological assemblages, this region is divided into three sub-regions: continental slope, shelf, and coastal regions (Ducklow et al., 2007). The shelf is 200 km wide with an average depth of 430 m, and quickly increases at the shelf break down to 3000 m at the continental slope. The WAP coastal region is divided into two straits (Gerlache and Bransfield) and Marguerite Bay. The Bransfield Strait is composed of three basins, which are separated by sills less than 1500 m deep. Gerlache Strait to the west of the Bransfield Strait has exchange with the strait limited to the upper 500 m, due to the presence of a sill ~500 m deep at the western end of the Bransfield Strait (Hofmann et al., 1996). The presence of the sills limits the exchange to surrounding waters to the north and west. On the strait's eastern end there is open exchange with the Weddell Sea.

The Southern Ocean Indian sector has a relatively straight coastline with a few bays, extending south of 67º S (Lützow-Holm Bay and Prydz Bay) (Fig. 1-C). One of the two major basins in this area is Enderby Abyssal Plain, which is located to the west of Kerguelen Plateau and to the north of Lützow-Holm Bay (Nicol and Raymond 2012).

1.2 Currents in the Southern Ocean

The Southern Ocean has the World's largest ocean current, the Antarctic Circumpolar Current (ACC). As the name indicates, this current flows continuously around the globe in an eastward direction and is driven by westerly winds. The ACC has three main components: the Subantarctic Front (SAF), the Polar Front (PF), and the Southern ACC Front (SACCF) (Orsi et al., 1995). The ACC extends from the sea surface to depths of 2000-4000 m and can be as wide as 2000 km.

The path of the ACC is mainly controlled by bathymetry; a northward shift in the path is observed around the Weddell Sea gyre, leading to a poleward shift bringing the ACC closer to the continent between 40º E and 160º E (Orsi et al., 1995). The second northward shift of the path is at the Ross Sea gyre, right after the ACC migrates towards the self-break. The Western Antarctic Peninsula region is unique due to its proximity to the ACC (Hofmann et al., 1996; Rintoul 2009).

The Antarctic Coastal Current (CC) flows westward adjacent to the Antarctic coastline, counter to the ACC. Although the ACC is circumpolar, the Antarctic Peninsula partially impedes its flow.

Circulation in the WAP region involves two main currents: the Antarctic Peninsula Coastal Current (APCC), first described by Moffat Varas (2007), forms during the ice-free seasons and flows south along Adelaide Island into Marguerite Bay (Moffat et al., 2008). The ACC flows northeast with its southern front almost parallel to the shelf break. Two cyclonic sub-gyres have been described for the shelf area of the WAP region (Smith et al., 1999). The proximity of the ACC and the presence of deep canyons connect the offshore and the coastal region, and also facilitate the exchange of nutrients and other water properties.

The CC and the Antarctic Slope Front merge at the northeast mouth of the Bransfield Strait. Thompson et al. (2009) showed drifter data indicating that water from the Weddell Sea is being transported into the Bransfield Strait via these two main currents. The coastal current flowing westward in Bransfield Strait may act as a transport mechanism to Marguerite Bay, and ultimately to offshore areas to the southwest. The drifter data from Thompson et al. (2009) also indicated the presence of a large standing eddy in the center of the Bransfield Strait that may act as a retention mechanism for zooplankton species like *S. thompsoni*.

In the Southern Ocean Indian sector, there are four circumpolar fronts, from north to south: Subantarctic Front (SAF), Polar Front (PF), Southern Antarctic Circumpolar Current Front (SACCF), and Southern Boundary of the Antarctic Circumpolar Current (SB-ACC) (Orsi et al., 1995; Nicol and Raymond 2012).

The circulation patterns in the Indian sector (30º E – 80º E) are determined by depressions and submarine canyons present in the shelf. However the eastward-flowing ACC located to the north and the more constrained westward-flowing CC are the main drivers of the circulation patterns in this region (Nicol and Raymond 2012). The water masses south of the ACC mostly come from the Australian Antarctic Basin by the westward-flowing Antarctic Slope Current (ASC). This current flows unbroken and forms the southern limb of the Weddell Gyre (Park et al., 2001). Intrusions of the ACC onto the shelf in this area were documented in 1987-1988; these bring in water properties and biological assemblages different than the ones found in the CC (Hunt et al., 2007).

1.3 Water Masses

There are three main water masses present in the WAP region. The relatively fresh and cold Antarctic Surface Water (AASW), with temperatures ranging from -1.8-1° C and salinities from 33.7 to 33 ppt (e.g., Smith et al., 1999), occupies the upper part of the water column. Due to heat loss in autumn and winter, the AASW is a nearly homogeneous water mass that extends down to about 250 m depth. Due to surface heating and sea-ice melting in summer, the AASW warms at the surface, while the salinity decreases. The mixed layer of the AASW in summer is mostly between 30-80 m thick.

In the summer the AASW has a distinguishable layer at 200 m with an intense temperature minimum, this water is also commonly known as Winter Water (WW). The WW is characterized by a low temperature of -1.9° C and low salinities as the result of ice melting in the summer in the upper 100-250 m of the water column. WW is formed when the mixed layer is cooled to temperatures found in the winter water. It remains there because the solar incidence during the austral summer is limited to the water above the mixed layer found in winter (Baum 2010).

The most prominent water in this region is the Circumpolar Deep Water (CDW) found between 200-700 m depth. This water mass is characterized by temperature higher than 1° C and salinity ranging between 34.6-34.73 ppt (Hofmann et al., 1996) and can be found in the Bransfield Strait as well. The presence of the southern boundary of the ACC along the outer WAP continental shelf produces subsurface, on-shelf intrusions of the CDW. Bransfield Strait has its own water mass called the Bransfield Strait Basin Bottom Water (BS), which is characterized by temperature below 0° C and salinity between 34.45-34.6 ppt (Stein and Heywood, 1994).

The main water masses in the Indian Sector are the Antarctic Surface Water (AASW), Circumpolar Deep Water (CDW), Modified Circumpolar Deep Water (MCDW), Shelf Water

(SW), Ice Shelf Water (ISW), and Antarctic Bottom Water (AABW) (Meijers et al., 2010). There is a north–south gradient, where water at stations very near the coast tends to be warmer and saltier, but becomes cooler and fresher to the north where AABW is observed, and then increases in temperature and salinity again near the ACC.

T-S diagrams of the WAP region only show the Antarctic Slope Front, as both the Southern Boundary and SACCF are found north of 65º S. The eastern-most extent of the Weddell Gyre occurs in the Indian Sector (Meijers et al., 2010).

1.4 Ecology and Distribution of Salpa thompsoni

Salpa thompsoni frequently forms dense blooms covering large areas of the Southern Ocean during the austral summer. Salps alternate between morphologically different asexual and sexual forms. The single asexual form (oozooid, solitary) is shaped like a cylinder and has a stolon that forms chains (clonal aggregate salps) that are eventually released into the water column when the conditions are favorable (i.e., austral spring and summer) (Loeb and Santora, 2012). The stage of the stolon development is used to characterize the life stage of the oozooids,

The sexual form (blastozooid, aggregate) has a more tapered shape that allows individual blastozooids to connect with one another in a chain. Each blastozooid in a recently-released chain has both a testis and an ovary; the ovary matures first and therefore fertilization can occur promptly after the chain has been released. The development of the embryo inside the blastozooid is used to characterize the maturity of each individual from Stage 0 (i.e., recently released chain from the oozooid) to Stage 4-5 (i.e., well-developed oozooid embryo) (Daponte et al., 2001).

Studies characterizing the different life stages of *S. thompsoni* have aided understanding of

population dynamics and life cycle adjustments in response to environmental variability (Chiba et al., 1999). This life cycle strategies facilitate rapid population increase when environmental conditions are advantageous and can often result in dense salp blooms (Alldredge and Madin, 1982).

The oceanography and zooplankton community composition have been documented for both the WAP region and the Indian sector of the Southern Ocean. However few studies have shown comparisons between temporal and spatial patterns of environmental conditions with respect to salp presence and absence. This work will describe similarities and differences in *S. thompsoni* habitat between regions and seasons with different environmental conditions, and will provide new insight into environmental conditions associated with salp distribution in the sampled Southern Ocean regions.

2. METHODS

2.1 Sampling Domain

Collections were made in the Western Antarctic Peninsula (WAP) region in the austral spring during November 10-25, 2011 onboard the R/V *L.M. Gould* (LMG 11-10). Twenty stations were sampled in the Bransfield Strait, continental shelf waters west of Anvers Island, shelf and shelf break waters west of the South Shetland Islands, and the deep ocean waters off the shelf (Fig. 1-A). For a detailed description of the scientific program of the cruise see the LMG 11-10 cruise report (Bucklin et al., 2011; http://data.bcodmo.org/LMG11-10/LMG11-10_Cruise_Report_06dec11.pdf).

The WAP region was also surveyed during the R/V *Polarstern* expedition (ANT XXVII-2) from January 9-30, 2011 (Fahrbach et al., 2011; http://epic.awi.de/29981/1/Fah2011b.pdf). Collections

were made at a total of 81 established standard stations on 10 parallel sections perpendicular to the coast, covering the area between the Bransfield Strait in the north and Adelaide Island to the south (Fig. 1-B). Each station included hauls with the Rectangular Midwater Trawl (RMT; Ocean Scientific International Ltd.; Baker et al., 1973; Roe and Shale 1979). Casts with CTD rosette and water sampler occurred at stations separated by 20 nm.

Collections were made in the Southern Ocean Indian sector during the austral summer in January 8-30, 2009 on board the TR/V Umitaka-Maru (UM08-09). This expedition departed from Cape Town (South Africa) and carried out oceanographic observations and sampling at 12 stations in the area south of the Enderby Abyssal plain and north of the Lützow-Holm Bay (34º-46º E and 59º-70º S) (Fig. 1-C).

2.2 Environmental Parameters

2.2.1. Temperature and salinity

During austral spring 2011 sampling in the WAP region, fluorescence, conductivity, and temperature were measured using Sea-Bird SBE CTD sensors at 20 stations (Fig. 1-A). The CTD was connected to a rosette with Niskin water bottles and sampling was carried out down to 1000 m, or 5-10 m above the sea floor in shallower waters. Water in the Niskin bottles was collected for measurements of nutrients, chlorophyll a, particulate organic carbon (POC), and particulate organic nitrogen (PON). The bottles were usually triggered at 5, 6, or 7 depths depending on the station depth (see LMG 11-10 cruise report; Bucklin et al., 2011). Nutrients, chlorophyll a, particulate organics (POC/PON) data were only available for the expedition carried out in the WAP region during the austral spring of 2011.

During austral summer 2011, a CTD cast was carried out at each of the 81 stations in the WAP region. The CTD was equipped with double sensors for temperature (SBE3plus) and conductivity (SBE4C). The CTD/water sampler consists of a SBE911plus CTD system combined with a carousel water sampler type SBE32 for 24 samplers and with 12 liter volume each. Salinity and temperature data collected during the CTD casts was obtained from the PANGAEA database (Rohardt et al., 2011) (http://www.pangaea.de/). Water samples were taken from the Niskin bottles (12 L) only at 15 stations for chlorophyll a and nutrient analysis. No particulate organics (POC/PON) were studied for this research expedition.

Sampling in the Indian Sector during austral summer 2009 included vertical profiles of water temperature and salinity taken at seven stations using a CTD (Model SBE911, Sea-Bird Electronics). Water samples from Niskin bottles were collected to analize nutrients and chlorophyll in the water column. Data for Particulate organics (POC/PON) were not measured.

2.2.2 Nutrients

In the WAP during austral spring 2011, a filtering capsule with a 0.2 μm filter was used to remove bacteria from water for nutrient analysis. Water was run through the capsule for 1 min. Three scintillation vials were filled with 10 ml and purged of any air. The samples were frozen immediately at -20° C for later analysis.

Nutrients were analyzed via an automated gas-segmented continuous-flow colorimetric method and measured with the Smartchem (Unity Scientific, Brookfield, CT) at the Department of Marine Sciences, University of Connecticut. Nutrient analysis was conducted following methods of the Environmental Protection Agency (EPA): EPA method 353.4 was used to determine the concentration of Nitrate (NOX) (Zhang et al., 1997); EPA method 366.0 was used to determine

the concentration of silicate (Zhang and Berberian 1997); and European Social Survey (ESS) method 310.1 (ascorbic color metric method) was used to determine the concentration of phosphate (USEPA, 1992)

2.2.3 Chlorophyll *a*

In situ measurement of chlorophyll *a* was carried out in the WAP during austral spring 2011. Water collected in the Niskin bottles during the CTD casts was used to measure chlorophyll *a* concentration in the water column. Water (500 ml) collected from the Niskin bottles was filtered through glass Whatman fiber filters (GFF) (20 μm) to collect phytoplankton; the filters were then frozen at -80^{o} C for later analysis at the Department of Marine Sciences, University of Connecticut. In total, 558 filters were analyzed and extracted with acetone following EPA method 445.0 (Arar and Collins 1997). The concentration of chlorophyll *a* was determined by measuring the fluorescence of the extracted pigments in a TD-700 Laboratory Fluorometer (Turner Designs, Sunnyvale, CA).

In order to get a continuous chlorophyll *a* profile for the water column, the chlorophyll *a* data obtained from the filters were used in a regression analysis to calibrate the fluorescence values collected during the CTD cast. The profile data were used in EasyKrig 3.0 (ftp://globec.whoi.edu/pub/software/kriging/easy_krig/) to create contour section plots.

The chlorophyll *a* values between the cruises were not comparable, given the different methods of collection and data analysis used. Some values were based on extracted chlorophyll *a*, while others were relative values based on the calibration of the instrument used.

2.2.3.2 Satellite chlorophyll *a* data

Surface chlorophyll *a* concentrations were derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua satellite images downloaded from the National Aeronautics and Space Administration (NASA) ocean color website and using the standard chlorophyll algorithm (Dierssen et al., 2000). Time-averaged monthly and 8-day composite chlorophyll images were obtained nominally at 4-km resolution for each sector and time period. White regions in the imagery denote areas that were obscured due to either clouds or sea ice for the given time period (see Fig. 5). Time-averaged imagery was utilized in this study due to the high amount of cloud cover common to Antarctic waters. Monthly composites do not represent the true monthly chlorophyll, but are calculated as pixel averages from only the limited number of days observable from the satellite. These data represent chlorophyll at the sea surface and are generally highly correlated with integrated chlorophyll concentrations within the euphotic zone (Dierssen et al., 2000). For each station, the chlorophyll value from the pixel closest to the sampling location was selected from the imagery.

2.2.4 Particulate organic carbon and nitrogen (POC, PON)

In the WAP during austral spring 2011, measurements were made of particulate organic carbon (POC) and nitrogen (PON) by filtering 0.5 L of water through pre-weighed and pre-combusted Whatman GF/F filters. The filters were frozen immediately and stored at - 80° C until analysis at the Department of Marine Sciences, University of Connecticut.

The instrument used for POC/PON analysis was a two elemental analyzer Fisons NA 1500 series (Costech Analytical Technologies Inc., USA). The operation principle was taken from the Costech analytical technologies website (http://www.costechanalytical.com/products/ecs4010.aspx). The method used for POC/PON

determination is based on EPA method 440 (Zimmermann et al., 1997), however sodium nitroprusside was used in place of the dichloroisocyanuric acid sodium salt.

2.3 Zooplankton Collection and Analyses

The austral spring 2011 cruise to the WAP region employed a Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS; Wiebe et al., 1985) with a mouth opening of 1-m^2 and nine 335-μm mesh nets. It was normally towed obliquely down to 1000 m and it sampled eight successive depth layers. At shallower stations, the net was towed down to as little as 350 m. Additionally, a 2.3 m^2 Isaacs-Kidd Midwater Trawl (IKMT; Isaacs and Kidd, 1953) with a 505 μm mesh net was employed. The net had a pentagonal mouth opening measuring 137.1 cm wide by 157.5 cm high, with two 81.3 cm wide lower panels. A deep oblique tow was lowered at ~20 m/min (wire out speed) to a depth of 175 m, and raised up at a wire-in speed of 10-15 m/min. A shallow tow followed at a depth of ~50 m at a wire-out speed of ~ 5 m/min and raised to the surface at the same speed (Cruise report LMG11-10; Bucklin et al., 2011).

During the 2011 austral summer cruise (ANT-XVII/2) in the WAP region, the Rectangular Midwater Trawl (RMT 1+8) equipped with a real-time time-depth-recorder (TDR) was used as standard gear to collect salp samples from the upper 200 m. The RMT 1+8 was equipped with two open/close net systems each having three nets. The first net system, RMT8, had an opening of 8 m² and a 4.5 mm mesh at the mouth, with the last 1.8 m of the nets with a mesh of 0.85 mm. The second system, RMT1, had an opening of 1 m² and a 300 μm mesh net (Baker et al., 1973). A double oblique net tow was carried out at all stations. The total time of the net haul from surface to bottom to surface was approximately 40 min (Cruise report ANT-XVII/2; Fahrbach et al., 2011). During the cruise, 1,488 samples collected from 40 different stations using the RMT

were examined for *S. thompsoni*. A total of 1,052 identified specimens was removed and prepared for molecular analysis.

In the Indian Sector during austral summer 2009, samples were collected also using a RMT 1+8. The RMT was towed obliquely in six different strata between the surface and 2,000 m (0–50, 50–100, 100–200, 200–500, 500–1000, 1000–2000 m). The ship's speed during the net tows was 1 m s^{-1}.

2.4 Handling of Salp Samples and Specimens

Salps species were identified under a dissecting scope and life stages were determined based on Daponte et al., (2001), using the developmental stage of the oozoid embryo (when present) in the blastozooid or the stage of the development of the blastozooid chain in the oozoid. Each salp was measured using the standard oral-atrial length (Daponte et al., 2001) with a caliper. No abundance data are available, but the occurrence of *S. thompsoni* at the stations was recorded as presence / absence. For molecular analysis, the stomachs and embryos were removed by dissection to avoid contamination from prey; the remaining tissue was flash-frozen in liquid nitrogen and stored at -80° C and shipped to the University of Connecticut for analysis.

2.5 Data Analysis

Multilevel comparisons were made for all the environmental data collected. Comparisons between areas (i.e., offshore and shelf) in the WAP region were made for each expedition. Comparisons between seasons (i.e., austral spring and summer) were done for the WAP region. Lastly, comparisons were made between the WAP region and the Indian Sector during austral summers of 2009 and 2011.

Temperature, salinity, chlorophyll *a* data collected with the Sea-Bird CTDs from selected stations (with corresponding latitude and longitude) during each expedition were used to generate contour plots using EasyKrig3.0 (http://globec.whoi.edu/software/kriging/easy_krig/easy_krig.html). Kriging was used to interpolate from the sparse original data. The kriging was done using the general exponential-Bessel model with nugget set to 0, sill < 1, length around 0.5, power > 1.5 and range ~ 0.5.

Salinity and temperature data were plotted using MATLAB R2013a (Mathworks, Natick, MA) to generate T/S plots and identify the signature of water masses present in the sampling domain. For the generation of T/S plots, CTD data from the top 1-4 m were ignored, due to the inaccuracy of the measurements; the salinity values below 5 m were found to be valid and consistent.

Environmental parameters such as salinity, temperature, nutrients, and POC were analyzed with a principal component analysis (PCA), using Euclidean distances in primer-E Ver. 6.0.1. PCA allows the characterization of habitats according to the eigenvectors by looking at the influence of each parameter on the principal components. Variables were standardized by dividing the values by the maximum value before carrying out the PCA (Clarke and Gorley, 2006)

Two different analyses were carried out: one using all the environmental data available for all stations, regardless of salp presence or absence. The second analysis was done using the environmental data only for those stations where *S. thompsoni* was collected.

3. RESULTS

3.1 Environmental Parameters

3.1.1 Temperature and salinity

The warmest waters (~2º C) were recorded in the WAP region during the 2011 austral summer. The coldest water (-1.8º C) was found in the Indian Sector in the stations close to the ice edge during the 2009 austral summer.

Temperature and salinity were measured at 24 stations during austral spring 2011 in the WAP and comparisons showed that temperature and salinity profiles in the offshore section shared characteristics observed at the stations going from southwest (Stn. 24) to the northeast (Stn. 11; Fig. 2). Although the more coastal Bransfield Strait Stns. 20 and 12 showed similar temperature and salinity profiles to the offshore section, the other stations showed the characteristic profiles of Bransfield Strait waters (Fig. 2 and 3), including cooler water below 150 m that is indicative of Bransfield Strait Basin Bottom Water (BS). At Stn. 15, there were intermediate or transitional conditions between the typical Bransfield Strait profile and the typical offshore profile, as observed at Stn. 12.

During austral summer 2011, the temperature profiles for both offshore and mid-shelf waters of the WAP showed warmer water at the surface compared to the profiles from Austral Spring. However, due to the presence of CDW, the offshore profiles showed warmer water below 200 m than the temperature profiles from the mid-shelf section (Fig. 3).

The profiles at northern stations in the Indian Sector sampled during austral summer 2009 were similar (Fig. 3). Stn. L3 is the northernmost of the stations; there were cooler temperatures at the surface, probably due to the influence of the ACC (Fig. 3). Due to their proximity to the ice edge, the southernmost Stns. L12 and L37 showed a decrease in the water column temperature with depth. Stn. L12 was the closest to the ice edge and had the freshest water at the surface. Stn. L37

was farther away from the ice edge, but still close to the coast, and might therefore have been influenced by the CC.

The environmental data from a subset of stations were used to generate section contour plots that allowed both the comparison of different regions within the WAP and Indian Sector, and seasonal changes within each area. Additionally, local water masses were identified in all the areas as described below.

The offshore transect in the WAP during austral spring 2011 had cold and relatively fresh Antarctic Surface Water from the surface down to 100 m. A zone of rapid increase of temperature and salinity occurred at the pycnocline between 100 and 200 m, and CDW occurred below the pycnocline, which was between 120-150 m (Fig. 3). In the offshore transect, the temperature and salinity profiles showed the warm deep water from the CDW (Fig. 3).

Offshore Stn. 8 showed strong similarities with Stn. 20 (south of the Bransfield Strait), suggesting that the presence of a channel connecting the offshore waters with the strait was allowing exchange between the two regions.

The CDW at the eastern part of the Bransfield Strait was identified as coming from the offshore ACC. The local BS was also evident as the less than -1° C water (Hofmann et al., 1996; Smith et al., 1999) (Fig. 2). The cold saline water at the surface was likely from the Weddell Sea east of the strait (Fig. 3).

The surface water variation in temperature and salinity in the coastal areas (Bransfield Strait and shelf) was likely due to the seasonal processes, such as ice melting and the development of the water column stratification caused by the increased solar radiance in the austral summer.

The contour plots show the presence of different water masses in transects of the different WAP regions and in the Indian Sector (Fig. 3). The offshore section through the WAP region shows the presence of CDW in both austral spring and summer (Fig. 3). During austral spring 2011, the AASW was composed mostly of WW up to the surface and its upper portion was warmed in summer by solar heating at the surface, thus overlaying a deeper layer with WW. Also in the Indian Sector the CDW was evident from 200 to 500 m in the Indian Sector during austral summer 2009. However, the CDW started to disappear in the south stations closest to the ice edge, where the deep water became cooler and fresher (Fig. 3-C). The coldest and saltiest water occurred at these stations, i.e., Stns. L12 and L37 (Fig. 3-C). According to Oshima (1996), the thickness of the WW layer at this location exhibits clear seasonal variations reaching its maximum in the austral fall (500 m) and minimum in the austral summer (350-400 m).

3.1.3 Chlorophyll *a*

The chlorophyll *a* concentration in austral spring 2011 in the WAP region differed greatly in the upper 100 m between the offshore and Branstfield Strait areas; with higher values at the more coastal stations in the Bransfield Strait (1.5 µg/L) compared to the offshore stations (1 µg/L). Below 100 m the concentrations were in both areas low ranging from 0.5 µg|L to 0.8 µg/L (Fig. 4 A, B). Satellite data showed higher chlorophyll *a* concentration in the WAP during austral summer 2011 compared to the spring, with the highest chlorophyll a values found in coastal and shelf areas (Fig. 5A-B).

In the Indian Sector during austral summer 2009, chlorophyll *a* concentration determined by analysis of collected phytoplankton and in situ CTD fluorescence profiles showed the same patterns, with higher chlorophyll concentration in the stations closer to the coast than in the

offshore stations. The contour plots show low values in the northernmost Stn. L3, increasing southward and reaching a maximum at the ice edge stations Stns. L12 and 37 (Fig. 4-C). The chlorophyll maximum was shallower (~30 m) at the southern Stn. L12 than the northern station L5 (~100 m). The satellite data clearly show the same trend of chlorophyll *a* concentrations, with high values at the coastal stations (Stns. L12 and L37), while the rest of the station offshore have lower concentrations (Fig. 5-C).

3.1.4 Nutrients

Nutrient (phosphate, nitrate, and silicate) concentrations were measured for 133 water samples collected in the WAP during austral spring 2011; only data from the upper 200 m (where salps were collected) are presented. Phosphate, nitrate, and silicate had the lowest concentrations in the upper part of the water column and increased with depth (Fig. 6). The nutrient profiles showed a steeper increase in concentration with depth in the offshore stations.

The Bransfield Strait had significantly higher values for silicate and nitrate (t-test; $p < 0.05$) than the offshore ones (Table 1). The largest difference was observed in the silicate concentrations between the two areas.

Phosphate, nitrate, and silicate were measured at 30 stations covering coastal, shelf, and offshore waters of the WAP during austral summer 2011. Only the nutrient data from the upper 200 m where salps were collected are presented. In contrast to the austral spring nutrient profiles, during this season the highest concentrations were at the surface and decreased with depth (Fig. 7). The nutrient profiles showed a steeper decrease with depth at stations that were on the shelf or in the shallower coastal region (i.e., stn. 188). In the offshore stations, nutrient concentrations remained the same with depth from 0-200 m.

Nutrients were measured at 13 stations in the Indian Sector during austral summer 2009. Nutrient profiles from the upper 200 m were similar to the ones measured in the WAP region during the austral spring of 2011. Nutrient concentrations were low at the surface and increased with depth (Fig. 8). Nitrate reached the highest concentration below 100 m in the offshore stations, whereas at the more coastal stations (Stns. L12 and L37), the highest value was at the bottom of the profile. Silicate concentrations increased with depth at offshore stations, whereas at coastal stations there was almost no change in concentration with depth. Phosphate was also lower in the upper water layers at all depths.

3.1.5. Particulate organic carbon and nitrogen

In general, particulate organic carbon (POC) decreased rapidly with depth in the WAP during austral spring 2011, with the steepest decrease in the Bransfield Strait region. The Bransfield Strait stations showed higher POC values at the surface compared to the offshore stations. Stns. 4 and 5 were located offshore and showed low POC values, which remained constant from the surface to 80 m, with a decrease in concentration below that depth (Fig. 9).

Measured C:N ratios were significantly higher in the Bransfield Strait compared to the offshore stations (Table 3).

3.2 Distribution and Life Stage Structure of Salpa thompsoni

In the austral spring of 2011, 206 samples collected by IKMT and MOCNESS from 8 different stations were examined for *S. thompsoni*. The salps were collected from 0-175 m using the IKMT; and from 0-50 m, 50-100 m, and 200-400 m using the MOCNESS (Fig. 10). In the WAP region in austral spring 2011 most of the salps were collected from 0-200 m. *Salpa thompsoni* was collected at the offshore stn. 4 (0-176 m) and seven stations in the Bransfield Strait (0-400

m) (Fig.11-A). Three stages of the solitary form (N=11) and early aggregate (stage 0, N=146) of *S. thompsoni* were found. The largest number of aggregate salps (N=81) was observed in the offshore region at Stn. 4, while only 65 aggregates were present in samples from all other stations in the Bransfield Strait. In contrast, more solitary salps (N=10) were found in six out of the seven stations in the Bransfield Strait and only one was collected offshore (Fig. 11-A).

During the austral summer of the same year, 1,488 samples collected by the RMT from 40 different stations were examined for *S. thompsoni*. A wide range of life stages was collected (aggregate 0-5; solitary 1-5) from 0-200 m.

Individuals of different life stages of the aggregate form were well-represented (Fig. 11-B). The oral-atrial length was higher for the same life stages for samples collected during the summer compared to the spring in the same area (Table 4). Male aggregates were collected only during summer and at this location, indicating a mature population that has gone through protogyny, (i.e., organisms begin life as females and change into males). Late stages of the solitary form were also collected. Solitary salps with embryos were only collected in the WAP region during the austral summer 2011. In terms of *S. thompsoni* life stage distribution in relation to season, we found that earlier stages (3-5) were present (but the oral-atrial length was the lowest (Fig. 12)) when compared to individuals collected later in the year in the same area (Table 4).

During austral summer 2009, a total of 117 salps was collected in the Indian Sector of the Southern Ocean during shallow and deep tows with the RMT at six stations covering an area from offshore (north) to the ice edge (south). Early life stages (0-4) of *S. thompsoni* were collected during the austral summer of 2009 in the Indian Sector. During the deep tows at depths between 1000-2000 m (Stn. L9) and 500-1000 m (Stn. L33), *S. thompsoni* was caught. This

agrees with the previous reports of this species being found not only at the surface, but also at greater depths (Ono et al., 2010; L.P. Madin (personal communication))

3.3 Multivariate analysis of environmental data and salp presence / absence

Principal Component Analysis (PCA) was done to examine the relationship between salp presence and environmental conditions. The PCA analysis of environmental data showed a main gradient on PC1, with two clusters clearly defined in the ordination plot representing data collected in the WAP region during austral spring 2011 and the Indian Sector during summer 2009 (Fig. 13). The same analysis showed a clear differentiation of the data collected in the WAP region during austral summer 2011.

In all, 80 percent of the variability was explained by PC1 and PC2 for all stations when temperature, salinity, and nutrient data were analyzed. The eigenvectors showed that the main factors influencing partitioning on PC1 were silicate, phosphate, and salinity, while temperature was the most influential for PC2 (Table 5). When the chlorophyll *a* data were included in the PCA analysis, only 63% of the variability was explained by PC1 and PC2. The eigenvectors showed that the main factors affecting partitioning of PC1 were silicate, phosphate, and chlorophyll a, while salinity was the most influential for PC2 (Table 6).

Analysis of environmental conditions at stations where salps were collected revealed that 83% of the variability was explained by PC1 and PC2. The eigenvectors obtained from the PCA analysis of the salp stations and temperature, salinity, and nutrient data again indicated that silicate, phosphate, and salinity were affecting partitioning of PC1, while nitrate was the main factor affecting PC2 (Table 6).

When chlorophyll *a* was included in the environmental data, the percentage of variability explained by PC1 and PC2 dropped to 74%. The eigenvector showed that the main factors affecting the partitioning of PC1 were the nutrients, while chlorophyll a was the most influential in the partitioning of PC2 (Table 7).

Whether all stations or just the salp stations were considered, both PCA analyses showed a clear differentiation between results from the WAP during austral summer and spring 2011. Initially, results from the WAP during austral spring 2011 and from the Indian Sector during austral summer 2009 were grouped together in the PCA (Fig. 14-B). The eigenvectors showed that the main factors influencing partitioning on PC1 were chlorophyll *a*, temperature, and nitrate, while silicate was the most influential for PC2 (Table 7). PCA analysis showed that 79% of the variability was explained by PC1 and PC2.

4. DISCUSSION

This study provides additional information useful for understanding and identifying the different environmental conditions in which *S. thompsoni* can be found, based on our field observations and collections from two Southern Ocean regions during two years (2009 and 2011) and both austral spring and summer. Characterization of the environment conditions associated with salp presence can help define habitat preferences of the species and thus provide a foundation for predictions of responses to changes in environmental conditions associated with climate change. In this study, we focus on salp presence / absence, as well as the relative frequencies of different life stages and forms, to provide an assessment of the state of the population at the time of sampling and the relationship between the measured parameters and season of the year.

The PCA analysis initially showed clustering of the austral spring 2011 (WAP) and austral summer 2009 (Indian Sector); these were clearly differentiated from the austral summer 2011 (WAP) (Fig. 13). When PCA analysis done for the WAP austral spring 2011 and Indian Sector austral summer 2009, showed clear differentiation was observed, putting into perspective the apparent similarity between WAP austral spring 2011 and Indian Sector austral summer 2009.

The eigenvectors of the PCA analysis of the WAP summer 2011 samples showed larger influence of nutrient concentration than of temperature or salinity. This is consistent with the summer environmental conditions, when the water column is warm but nutrient-depleted, thus having an effect on phytoplankton concentration and also affecting salp presence / absence. In contrast, the Indian Sector austral summer 2009 and WAP spring 2011 conditions are more affected by salinity; the obvious explanation for this is the proximity of the austral summer collections to the ice edge, where there was fresher water as a result of melting of the ice. Salp presence was not shown to be affected by nutrients; the melting of the ice and consequent release of nutrients into the surrounding waters clearly provided an environment rich in nutrients.

Chlorophyll *a* concentrations explained only a small portion of the observed variation, likely due to the measurement as a bulk property. Since salp grazing is strongly affected by phytoplankton cell size, and their preference for smaller cell size is well established, characterization of phytoplankton size distribution will be important to understand the relationship between chlorophyll and salp distribution. In contrast, silicate explained a significant fraction of the variation; this can be understood as a direct indicator of diatoms, due to the utilization of this element in their valves. Silicate might be a useful indirect indicator and predictor of salp distribution.

The PCA analysis of austral summer 2009 and austral spring 2011 data showed clustering of the samples along the chlorophyll *a* and temperature vectors for the spring 2011 collection. This may indicate the salps were in poor condition after overwintering. However, the samples from the austral summer 2009 were distributed more along the salinity and silicate vectors. This can be explained by the proximity to the ice edge and the resultant reduction in salinity due to melting ice, which favors *S. thompsoni.* The silicate measured in the water can be related to a diatom-rich phytoplankton community at the ice edge (Lizotte 2001).

Life stage determination depended on the integrity of the embryo inside each individual; the duration of the tow and the amount of zooplankton collected in each tow had an effect on integrity of the embryo. Longer tows or tows with high amounts of zooplankton will result in blastoozooids ripped apart, and embryos ripped from the inside of the blastoozooid due to turbulence or shearing with other plankton in the net.

During the spring 2011 expedition in the WAP region, the majority of the salps were found in the Bransfield Strait area. It has been suggested that the bathymetry and currents in the Bransfield Strait results in a central gyre that could act as a retentive mechanisms for zooplankton communities in this area (Thompson et al., 2009). During the spring, we found mature solitary forms that were starting to release chains of aggregates into the water column. During this time, the only stage collected was Stage 0 of the aggregate form. These observations agree with the proposition of the Bransfield Strait region as a retentive cell for zooplankton in the WAP (Thompson et al., 2009). Therefore an overwintering population of *S. thompsoni* might be retained in the Bransfield Strait. With the coming of spring, these populations may provide the basis for the salp blooms that are characteristic of this area later in the summer.

During summer in the WAP region, a wide range of life stages was found, suggesting that the aggregate individuals reached maturity (and larger sizes). It has been suggested that their vertical migration in association with CDW could allow southward transport (Loeb 2007) and populate other areas of the WAP region and form blooms throughout the region.

The salps collected in the Indian Sector of the southern ocean during the austral summer 2009 were more abundant along the eastern side of the cruise track. This area is closer to the coast and therefore shallower than the western side of the cruise track. The largest number of salps was collected at Stns. L33 and L37, with low (0.5 µg /L) and high (3 µg /L) chlorophyll a respectively. Stn. L37, located next to ice edge, had low temperature and high chlorophyll values, as usual for this type of environment (Ohshima et al., 1996). Early stages of aggregates and only one solitary were collected during this expedition. In order to understand the population dynamics of this species in this area multiyear studies are needed.

The Bransfield Strait has been reported as a nursery area for Antarctic krill (Loeb et al., 2010). Since early stages of *S. thompsoni* were collected in this area during the austral spring, which is spawning season for salps (Loeb et al., 2010; Loeb and Santora 2012), it is possible that the Bransfield Strait serves as a nursery area in the spring and that the currents in this area (Sangrà et al., 2011) serve as a retentive mechanism during the austral winter. As described above, these retained populations may form the basis for salp blooms when environmental conditions are favorable during the austral summer. The presence of only early aggregate (0) and early solitary stages during spring in the WAP region suggests that immature *S. thompsoni* may require more suitable environmental conditions to form dense blooms. On the other hand, the wide range of aggregate stages, aggregate males, late stages of the solitary form, and solitary embryos collected

in the WAP during austral summer 2011 indicates that a more mature population can generate from an overwintering population of solitary forms.

The differences in presence/absence of salps and their life stage distributions between seasons in the WAP region can be explained primarily by the timing of the collections. Fewer salps were collected in the spring prior to the arrival of favorable conditions, compared to the presence of salps and a wider range of life stages in the summer, after a couple months of favorable conditions. However as in previous studies (Atkinson et al., 2004; Lee 2010; Saba et al., 2014), winter sea ice cover is an important consideration for understanding salp distributions. During the winter previous to the austral spring 2011, the WAP region had higher sea ice concentration (14.5 M km^2) (Fig. 15-A). This will translate into unsuitable conditions for salps. In contrast, during the previous winter in 2010, the WAP region had lower sea ice concentration (14.3 M km^2) (Fig. 15-B). On the other hand the Indian Sector showed high ice coverage for the previous winter in 2008 (Fig. 15-C). The difference in ice coverage is most evident in the WAP region, where sections were covered by ice during austral winter 2011 but not during austral winter 2010. More studies comparing salp presence / absence and life stage distribution are needed during summer in different regions and under variable environmental conditions in order to understand the habitat preferences of *S. thompsoni.* Multi-year studies are crucial to understand the distribution of the Southern Ocean salp, *S. thompsoni* and how the species might be affected by changes in the marine ecosystem as a result of climate change.

5. CONCLUSIONS

This study took place within the well-studied waters of the WAP region (Hofmann et al. 1996; Thompson et al., 2009) and the Indian Sector (Sparrow et al., 1996; Ohshima et al., 2000). The

environmental data presented here showed the seasonal differences between spring and summer in the WAP region, including the characteristic seasonal modification of these water masses. Water mass features and distribution in the sampled areas were consistent with previous studies describing the presence and seasonality of this water masses in these regions.

Salps collected during the austral spring of 2011 were associated with Winter Water (WW) and Circumpolar Deep Water (CDW) in the offshore station and with WW and Bransfield Strait Basin Bottom Water (BS) in the Bransfield Strait stations. During austral summer 2011 in the WAP region, salp presence was associated with WW and CDW on the offshore and shelf stations. Similarly during the austral summer 2009 in the Indian Sector, salps were present at stations where WW and CDW water masses occurred in the offshore station, while more coastal stations were influenced by the Coastal Current (CC).

Based on long-term studies, it has been suggested that nutrients might also have a role affecting zooplankton distribution with the ongoing climate change (Lee 2010). Nutrient distribution was similar in all stations, regions and seasons. Our PCA analysis for austral spring and summer of 2011 in the WAP and austral summer 2009 in the Indian Sector showed that the main factors influencing *S. thompsoni* distribution were silicate (PC1) and temperature (PC2). However when chlorophyll *a* was included PCA analysis for austral spring 2011 in the WAP and austral summer 2009 in the Indian Sector, the main factors influencing distribution of this species were chlorophyll *a* and temperature (PC1) and silicate (PC2).

PCA analysis included in this work contributes to an understanding of *S. thompsoni's* distribution in relation to different environmental parameters (including nutrients).

This work also highlights the differences in life stage distribution in the WAP region during the

austral spring and austral summer as well as the difference in *S.thompsoni* life stage distribution during the austral summer in two different regions of the Southern Ocean, the WAP and Indian Sector.

Our field observations and collections from two Southern Ocean regions during two years (2009 and 2011) and both austral spring and summer provide additional understanding and identification of the different environmental conditions in which *S. thompsoni* can be found. Characterization of the environmental conditions associated with salp presence can help understand habitat preferences of the species and thus provide a foundation for predictions of responses to changes in environmental conditions associated with climate change.

Figure Legends:

Figure 1. Map of sampling domains and cruise tracks. (A) WAP region in austral spring (2011); (B) WAP region in the austral summer (2011); and (C) Indian Sector in austral summer (2009).

Figure 2. Temperature-Salinity diagrams of data collected in austral spring and summer (2011) in the WAP region (A, B) and during summer (2009) in the Indian Sector (C) of the Southern Ocean. Boxes represent different water masses: Circumpolar Deep Water (CDW), Winter Water (WW), Antarctic Surface Water (AASW) and Bransfield Strait Basin Bottom Water (BS).

Figure 3. Temperature contour plots generated by kriging data collected in the WAP region (2011) and in the Indian Sector (2009). Stations are shown in a southwest-to-northeast direction in the WAP region, (i.e., farthest station to the right is the easternmost station for that section) and north-to-south direction in the Indian Sector. Water masses in bold, Winter Water (WW), Circumpolar Deep Water (CDW) and Bransfield Strait Basin Bottom Water (BS). White vertical lines represent the depth of the CTD cast and therefore the range of depths where the data was collected.

Figure 4. Chlorophyll *a* contour plots generated by kriging chl *a* data collected in the WAP region (2011) and relative chl *a* data from the Indian Sector (2009). Chl *a* scales are different for

the WAP region and the Indian Sector. Stations are shown in a west-to-east direction, (i.e., farthest station to the right is the easternmost station for that section) in the WAP region and north-to- south direction in the Indian Sector. Red boxes indicate were *S. thomposni* was collected.

Figure 5. Chlorophyll *a* at the sea surface derived from the MODIS Aqua satellite downloaded from the National Aeronautics and Space Administration (NASA) ocean color website and using the standard chlorophyll algorithm. Time-averaged monthly and 8-day composite chlorophyll images nominally at 4-km resolution for each sector and time period. (A) WAP region during austral spring (2011); (B) WAP region during austral summer (2011); and (C) Indian Sector during austral summer (2009).

Figure 6. Phosphate (green), Nitrate (blue), and Silicate (red) profiles for the offshore and the Bransfield Strait stations sampled in the WAP region during austral spring (2011).

Figure 7. Phosphate (green), Nitrate (blue), and Silicate (red) profiles for the offshore and the Bransfield Strait stations sampled in the WAP region during austral summer (2011).

Figure 8. Phosphate (green), Nitrate (blue), and Silicate (red) profiles for the Indian Sector stations sampled during austral summer (2009).

Figure 9. Total Organic Carbon (TOC) measured in ug/L. blue profiles indicate TOC measurements made at offshore stations; red profiles refer to TOC measured at stations located in the Bransfield Strait.

Figure 10. *Salpa thompsoni* collection depths. (A) Salps collected with the IKMT and MOCNESS (moc) in the offshore (blue) and Bransfield Strait (orange) sections of the WAP region during austral spring (2011). (B) Salps collected during a shallow (s) or deep (d) tow of the RMT in the Indian Sector during austral summer (2009).

Figure 11. Life stages of *Salpa thompsoni* collected in: (A) WAP region during austral spring (2011); (B) WAP region during austral summer (2011); and (C) Indian sector during austral summer (2009). Different life stages are represented by the different colors (see legend).

Figure 12. Frequencies of individuals with different oral – atrial (OA) length (mm) for each aggregate life stage and solitary stage of *S. thompsoni* collected during the three expeditions.

Figure 13. Principal component analysis (PCA) of environmental data from stations surveyed in the WAP region during the austral spring (2011) (red) and the austral summer (2011) (blue); and in the Indian sector, surveyed during the Austral summer 2009 (green).

Figure 14. Principal component analysis (PCA) of environmental data from stations where *S. thompsoni* was collected in the WAP region during the austral srping-2011 (red) and austral summer (2011) (blue), and in the Indian sector surveyed during austral summer (2009) (green). (A) PCA analysis of temperature, salinity and nutrients; (B) PCA analysis of environmental parameters including chl *a of* two different seasons and regions: austral spring (2011) in the WAP and austral summer in the Indian Sector (2009).

Figure 15. Monthly Antarctic sea ice concentration images for 2011 (A, B) and 2009 (C). The red square highlights the WAP region for comparison of sea ice concentration between the winters of 2010 and 2011. Image courtesy of the National Snow and Ice Data Center, University of Colorado, Boulder (http://nsidc.org/data/seaice_index/).

Tables:

Table 1. Chlorophyll *a* and particulate organic carbon (POC) measured in the WAP region during austral summer (2011).

Region	Station	Depth (m)	Chl *a* (µg/L)	POC (µg/L)
Shelf	125	45	0.6	0.05
	160	30	1.2	0.13
	172	20	1.2	0.21
	175	25	28.7	1.01
	177	25	16.9	0.51
	178	40	1.3	0.07
	188	25	2.0	0.15
	191	20	10.3	0.61
	194	25	6.6	0.24
Offshore	133	55	0.1	0.06
	134	80	0.2	0.04
	167	75	0.4	0.06
	182	50	0.8	0.14
	184	40	0.3	0.10
	198	65	0.3	0.06

Table 2. Phosphate, Silicate, and Nitrate Two-sample t-test (assuming unequal variances) for the offshore and the Bransfield Strait stations sampled in the WAP region during austral spring (2011). P-values are shown in the second column.

T-test: Two-Sample (Assuming Unequal Variances)					
Nutrient	**P-value**	**Region**	**Mean**	**Min**	**Max**
Phosphate	4.36E-03	Bransfield Strait	1.72	1.60	1.85
		Offshore	1.81	1.60	2.15
Silicate	2.078E-05	Bransfield Strait	60.26	52.78	71.03
		Offshore	48.91	26.86	67.35
Nitrate	8.24E-03	Bransfield Strait	32.44	29.33	38.22
		Offshore	30.48	25.41	40.76

Table 3. Results of Two-sample t-test (assuming unequal variances) for C:N ratios measured at offshore and Bransfield Strait stations in the WAP region during austral spring (2011).

t-test: Two-Sample Assuming Unequal Variances		**Mean**	**Min**	**Max**
C:N ratios	**Bransfield**			
P-value=1.69E-03	**Strait**	3.81	**0**	**6.6**
	Offshore	2.59	**0**	**4.5**

Table 4. Oral – atrial (OA) length of the different life stages of the sexual and asexual form of *S. thompsoni* collected during the three expeditions.

		Number of individuals (OA length in mm)		
Aggregate Stages	**Total**	**WAP Spring – 2011**	**WAP Summer – 2011**	**Indian Sector Summer – 2009**
0	205	146 (5.8- 13.4)	38 (6.6-15.3)	21 (6.3-10.1)
1	157		138 (7.6 – 28.4)	19 (12.5- 27.9)
2	267		221 (11.5- 36.3)	46 (15.5-31.8)
3	223		198 (13.1-39.9)	25 (24.4-33)
4	86		85 (17.4 – 44.5)	1 (28.1)
Male	170		170 (16.3 – 52.8)	
Solitary Stages				
3	1	1 (56)		
4	9	8 (44.4 -78.3)	1 (80.9)	
5	10	3 (82.7 – 137.6)	7 (119.4 – 142.7)	
Solitary embryo	6		6 (13.6 – 25.3)	

Table 5. Eigenvectors from Principal Component Analysis of environmental data (temperature, salinity and nutrients) from three different expeditions at stations where salps were (presence) and were not (absence) collected.

All stations			
Variables	PC1	PC2	PC3
Temperature	-0.361	**-0.687**	-0.006
Salinity	**0.445**	0.406	0.277
Phosphate	**-0.465**	0.269	0.704
Silicate	**-0.538**	0.183	0.111
Nitrate	-0.407	0.506	-0.644

Table 6. Eigenvectors from Principal Component Analysis of environmental data (temperature, salinity and nutrients) available from the three expeditions at stations where salps were (presence) and were not (absence) collected.

Salp stations			
Variables	PC1	PC2	PC3
Temperature	-0.414	-0.474	0.763
Salinity	**0.438**	0.376	0.578
Phosphate	**-0.475**	0.042	-0.123
Silicate	**-0.518**	0.082	-0.158
Nitrate	-0.377	**0.791**	0.206

Table 7. Eigenvectors from Principal Component Analysis on environmental data (temperature, salinity and nutrients) from the WAP in austral spring (2011) and Indian Sector in austral summer (2009).

***Salp stations* – Spring 2011 and Summer 2009**			
Variable	PC1	PC2	PC3
Chlorophyll *a*	**-0.589**	0.019	0.112
Temperature	**-0.58**1	-0.067	0.003
Salinity	-0.046	**-0.565**	**0.555**
Phosphate	0.241	**-0.494**	**-0.649**
Silicate	0.092	**0.609**	0.202
Nitrate	**-0.497**	-0.245	**-0.465**

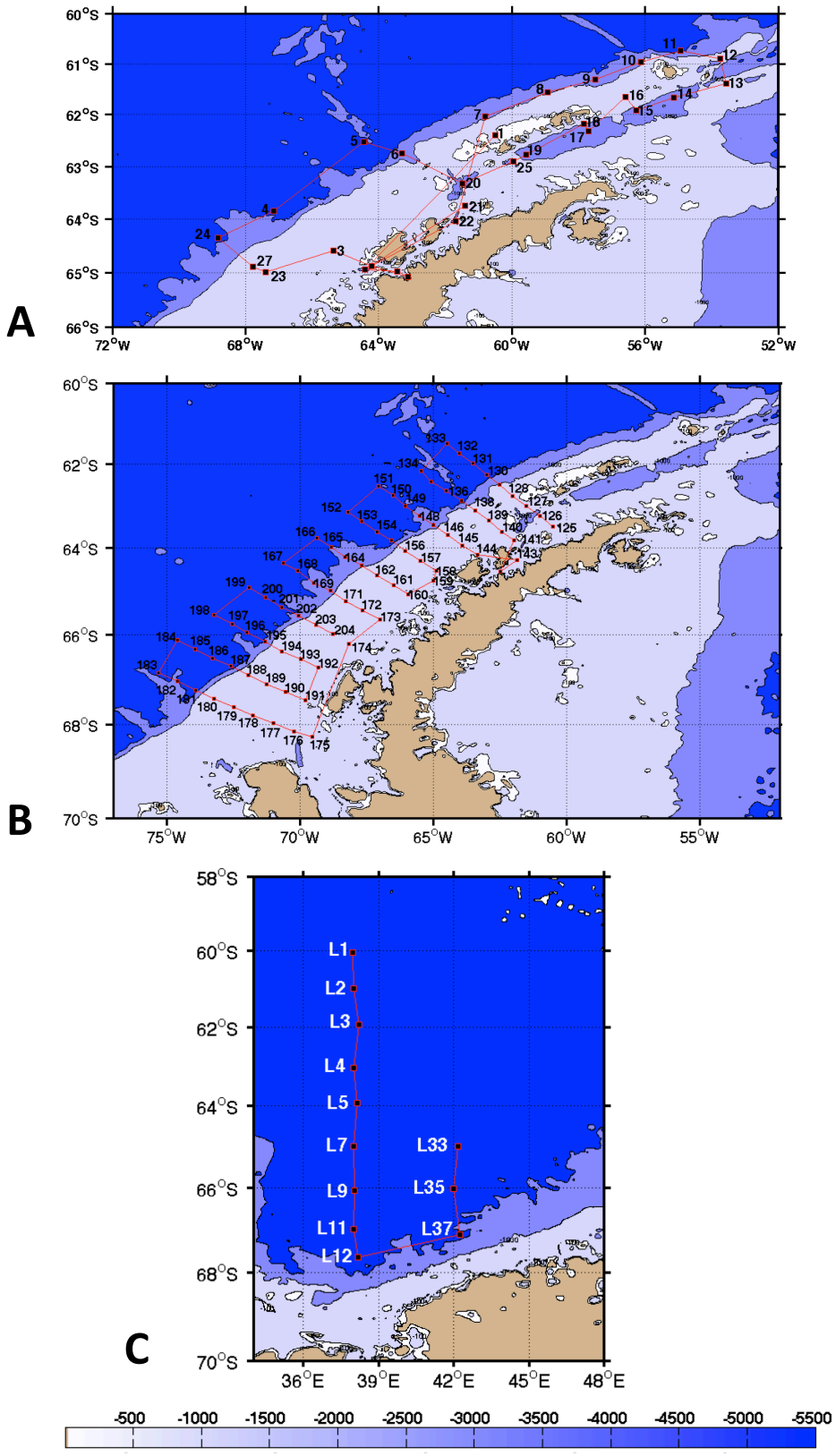


Figure 1. Map of sampling domains and cruise tracks

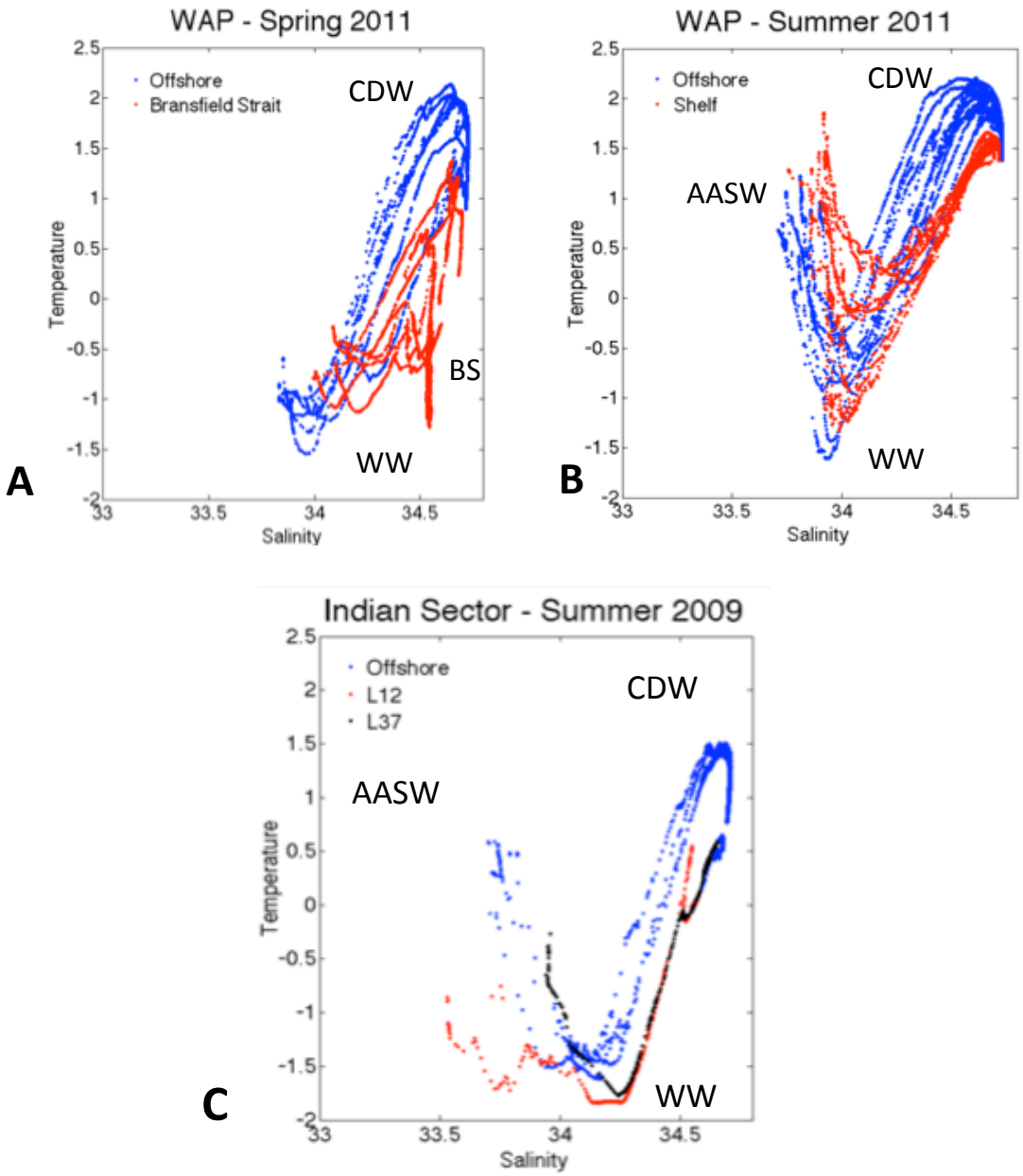


Figure 2. Temperature-Salinity diagrams

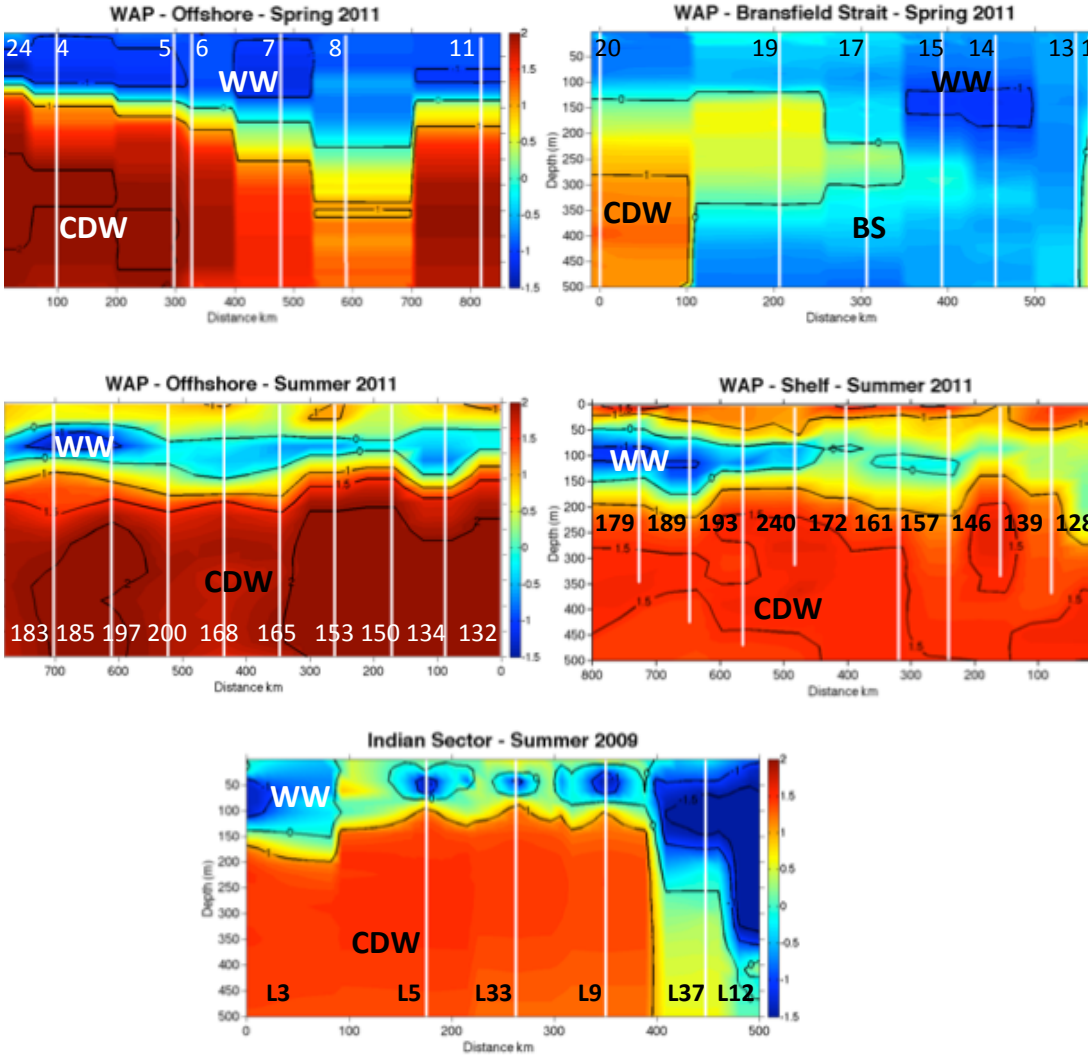


Figure 3. Temperature contour plots generated by kriging

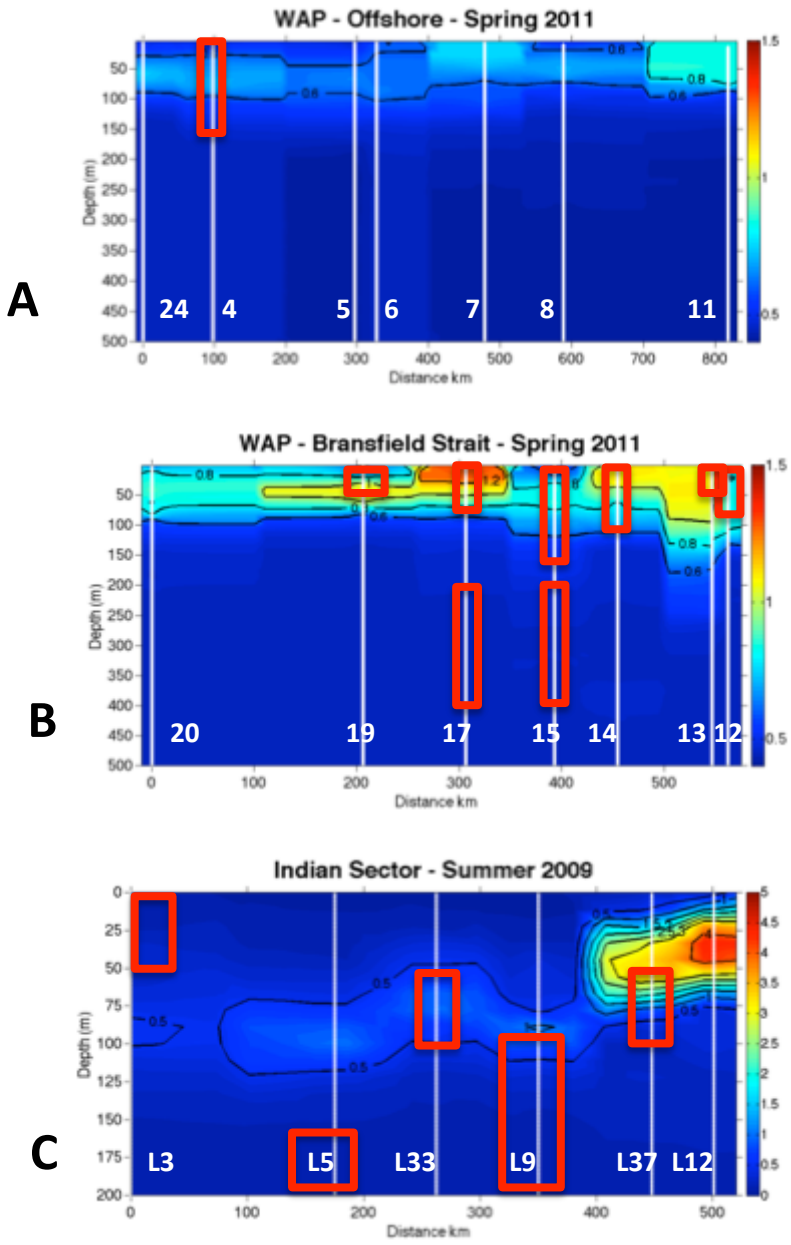


Figure 4. Chlorophyll *a* contour plots generated by kriging chl *a* data

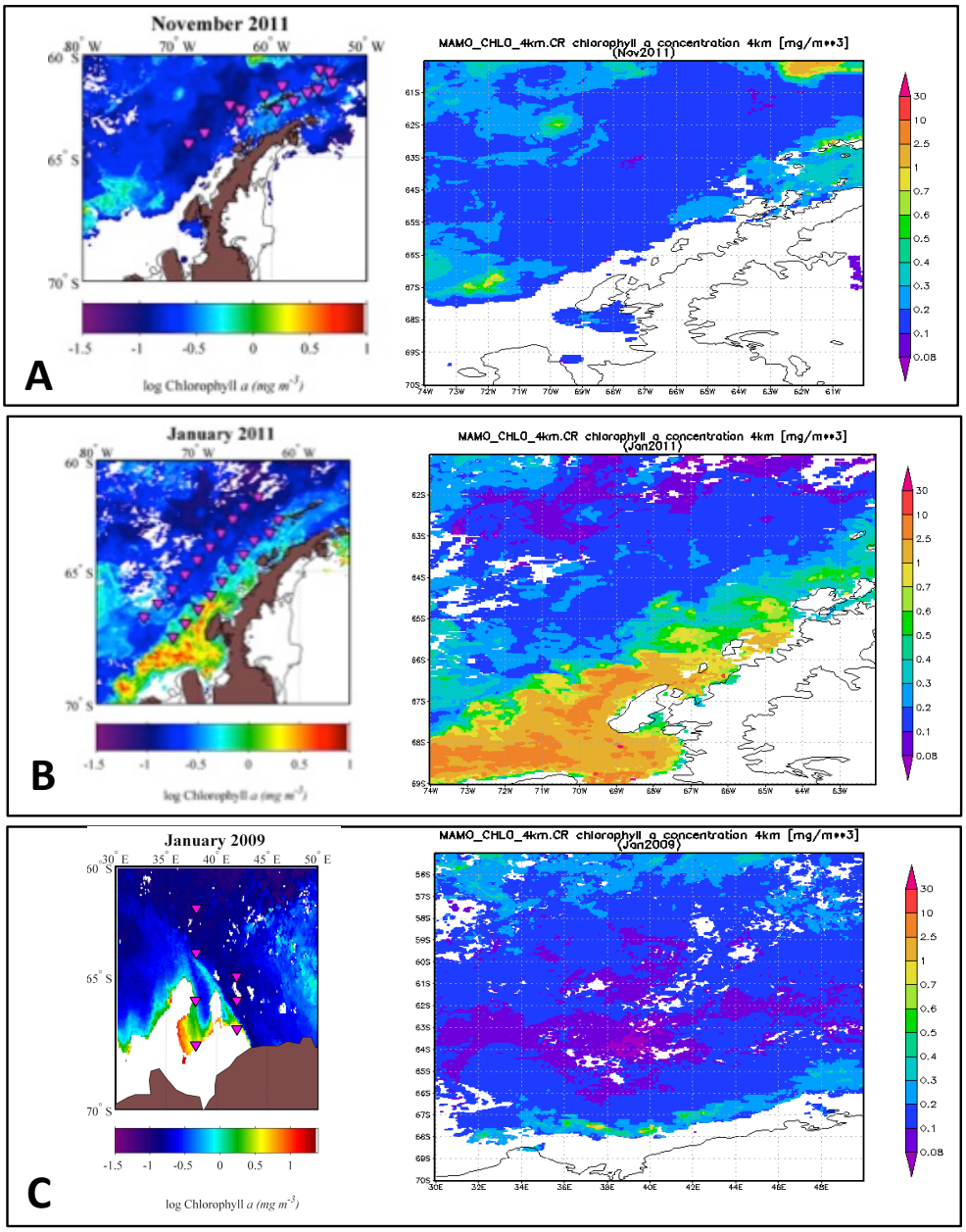


Figure 5. Chlorophyll *a* at the sea surface derived from the MODIS Aqua

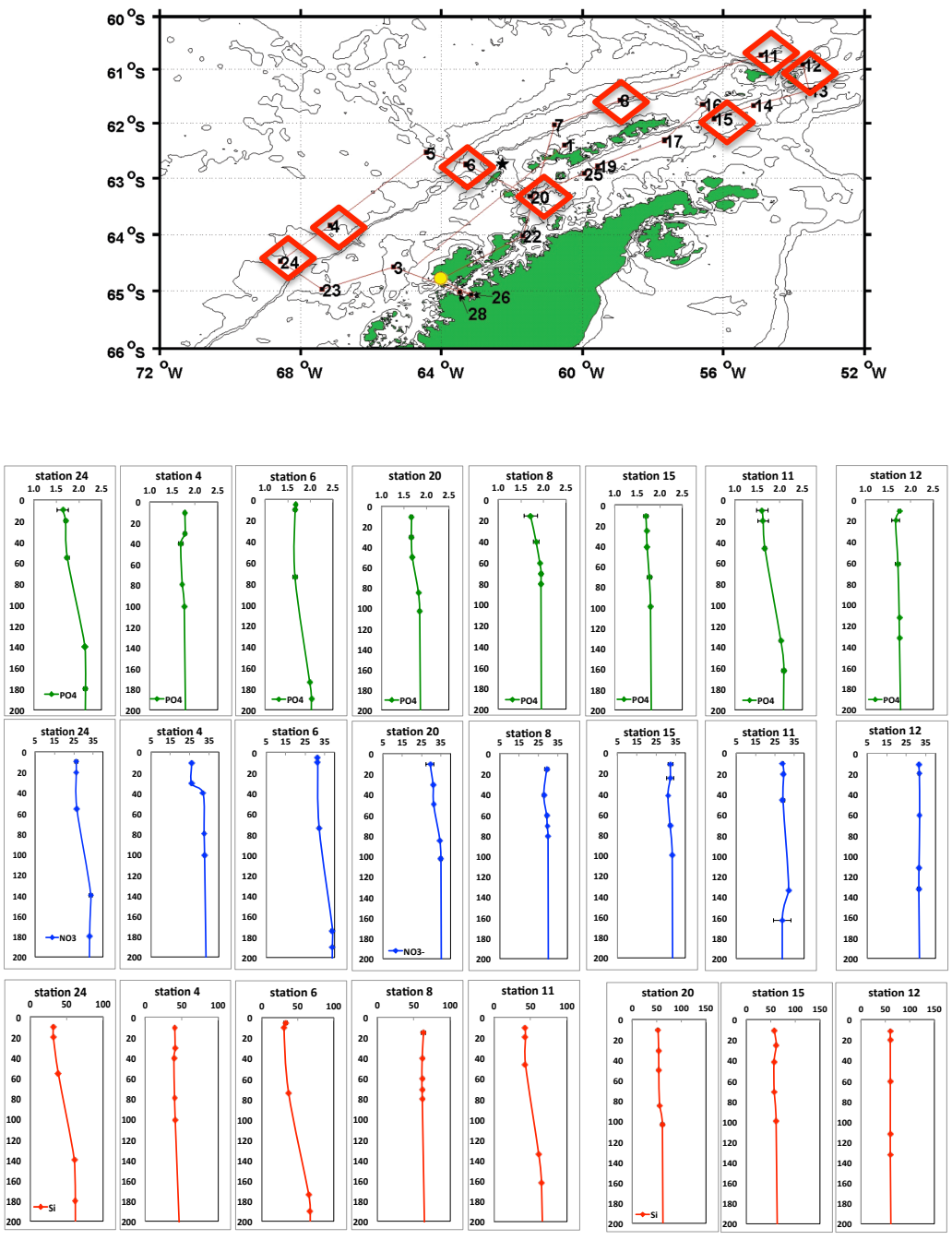


Figure 6. Phosphate (green), Nitrate (blue), and Silicate (red) profiles

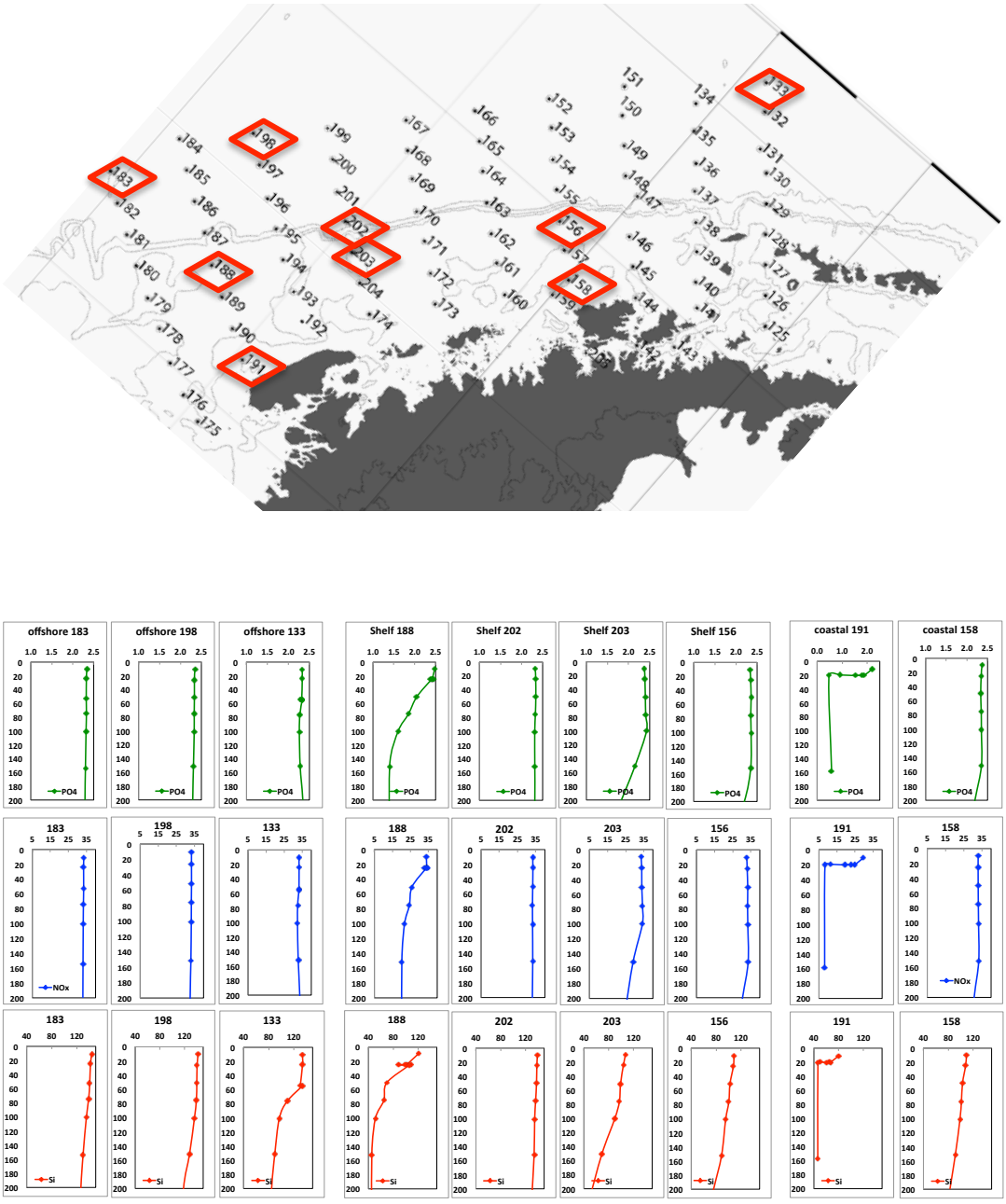


Figure 7. Phosphate (green), Nitrate (blue), and Silicate (red) profiles

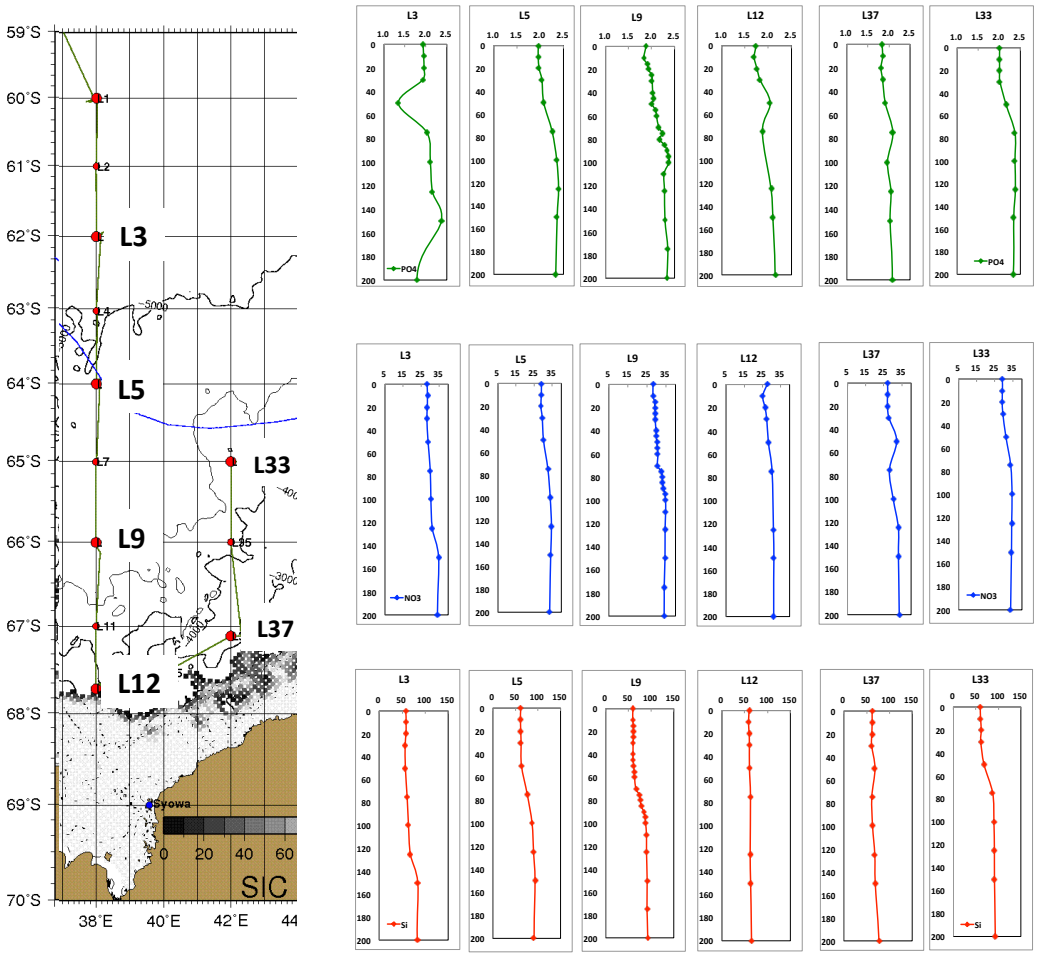


Figure 8. Phosphate (green), Nitrate (blue), and Silicate (red) profiles

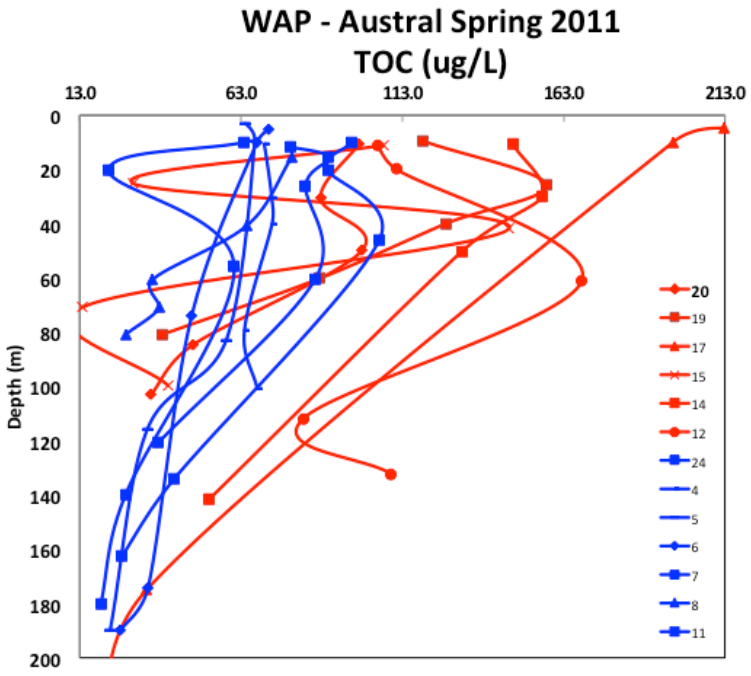
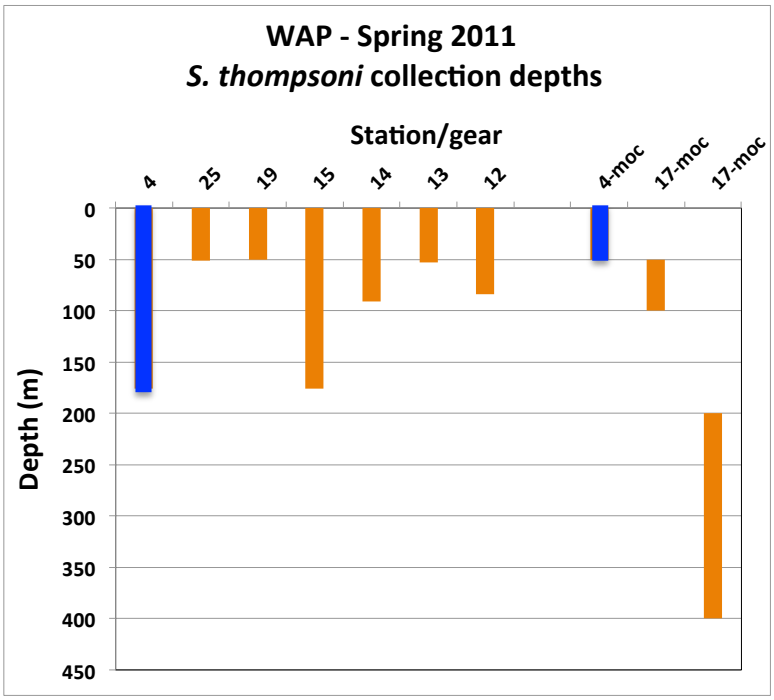


Figure 9. Total Organic Carbon (TOC) measured in ug/L.

A



B

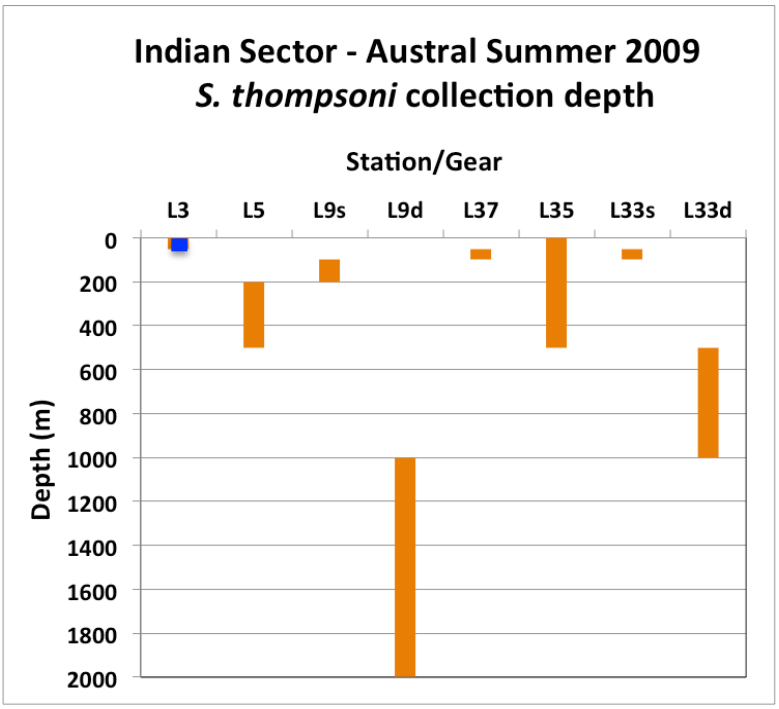


Figure 10. *Salpa thompsoni* collection depths..

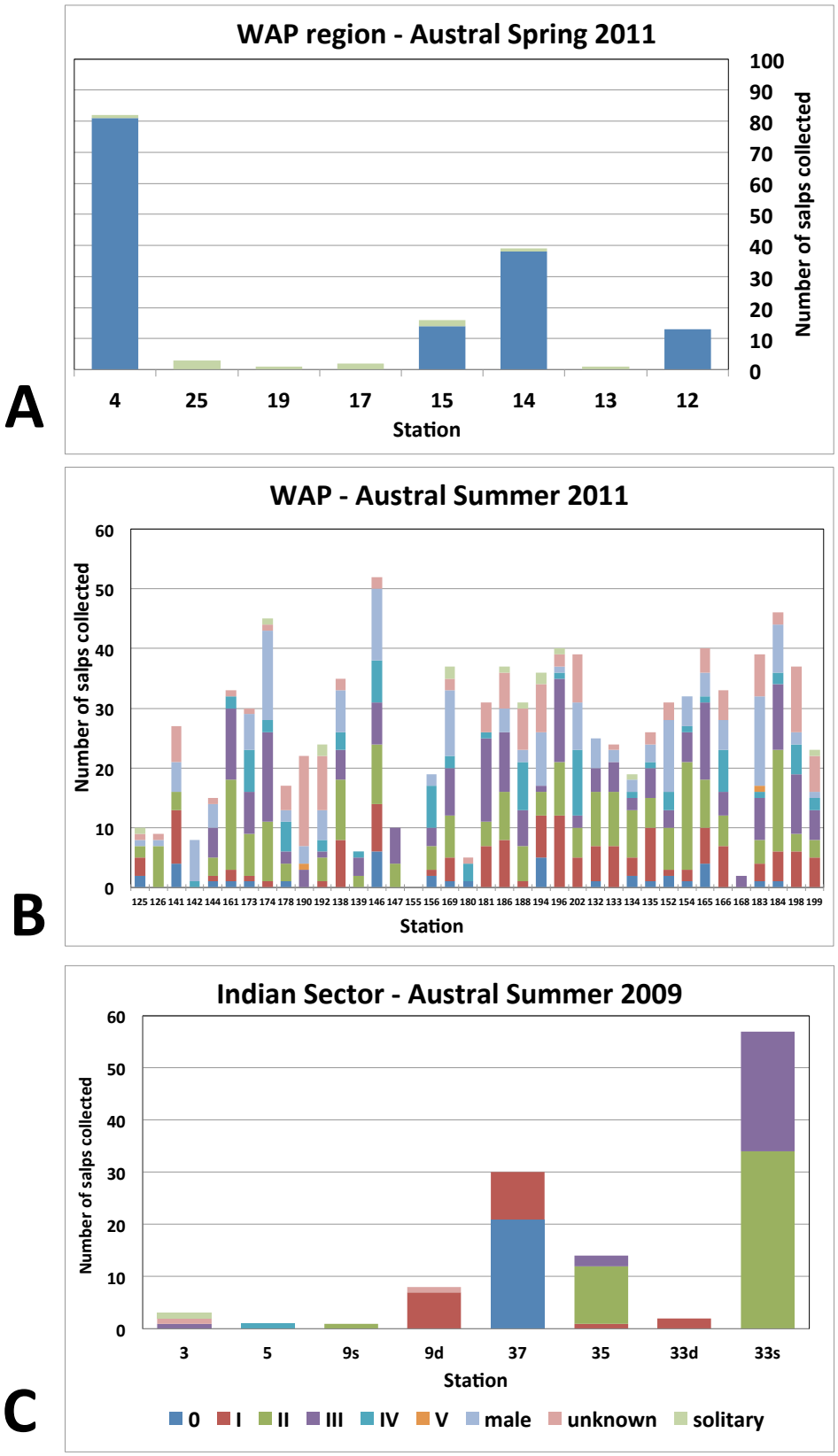
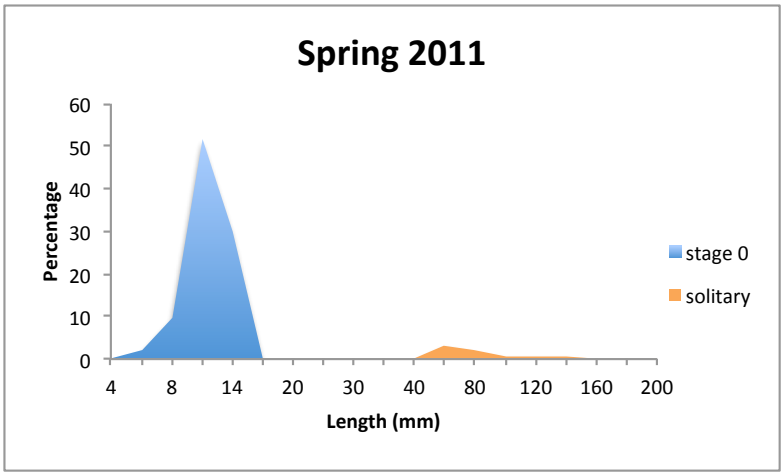
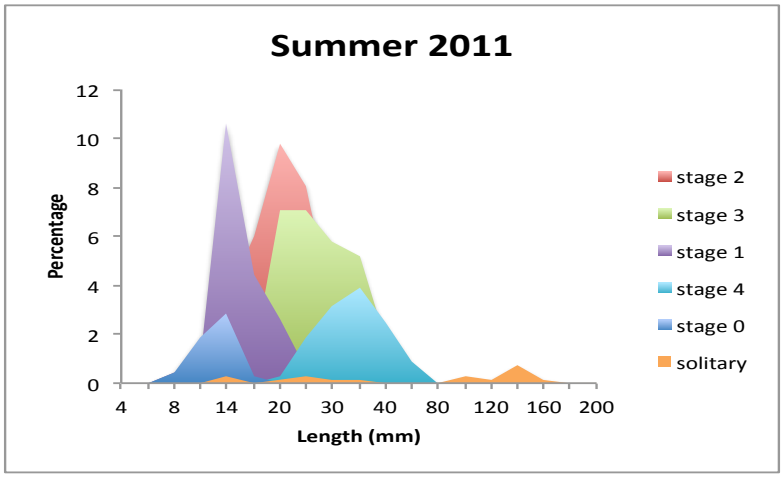


Figure 11. Life stages of *Salpa thompsoni* collected





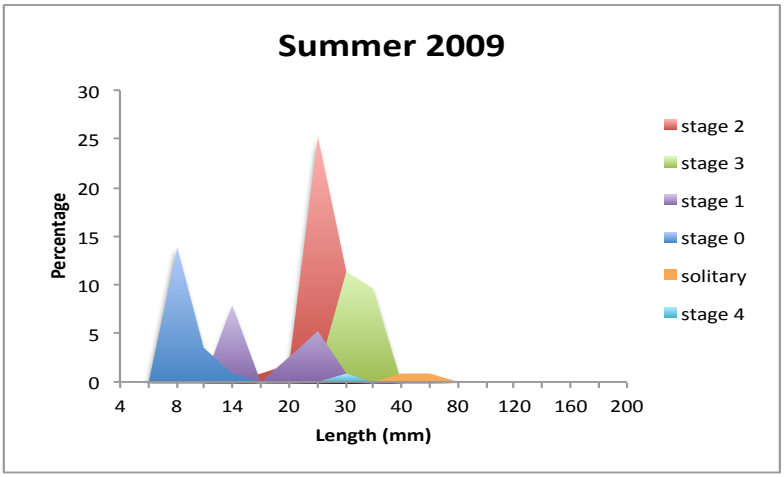


Figure 12. Frequencies of individuals with different oral – atrial (OA) length (mm)

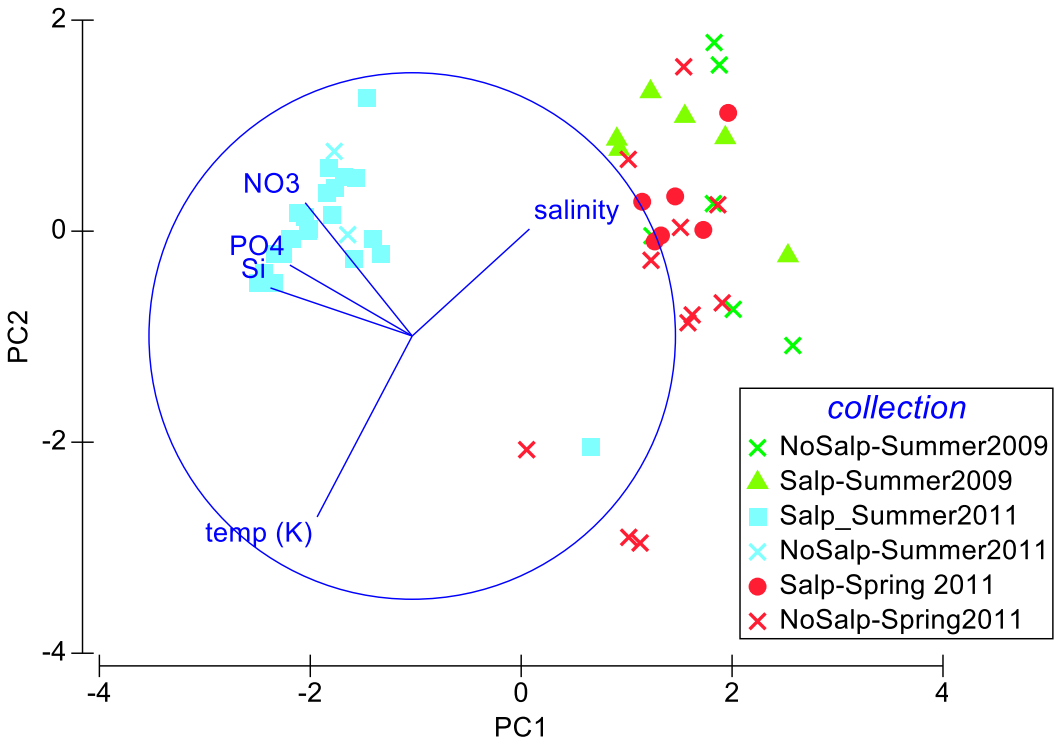


Figure 13. Principal component analysis (PCA) of environmental data.

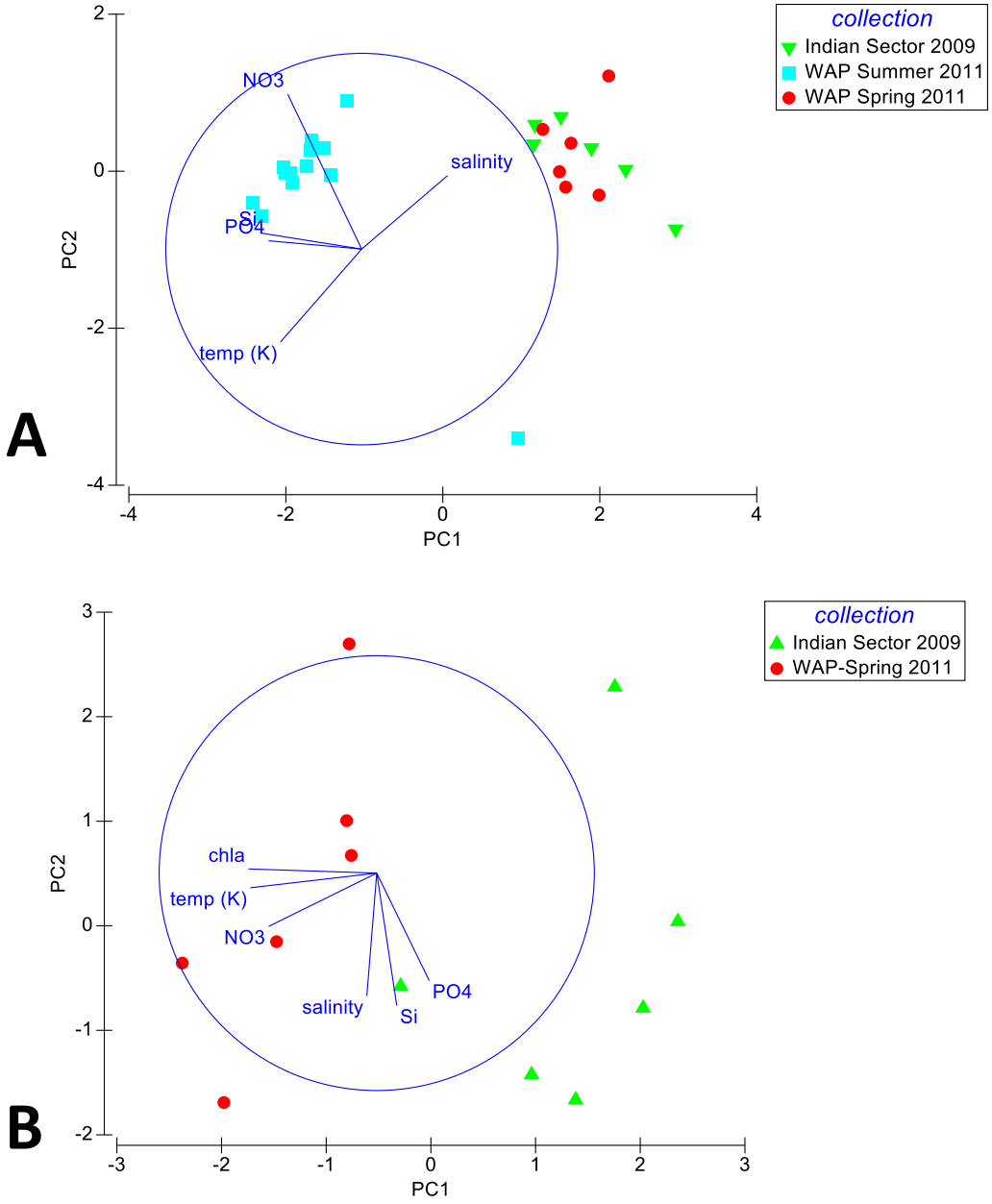


Figure 14. Principal component analysis (PCA) of environmental data

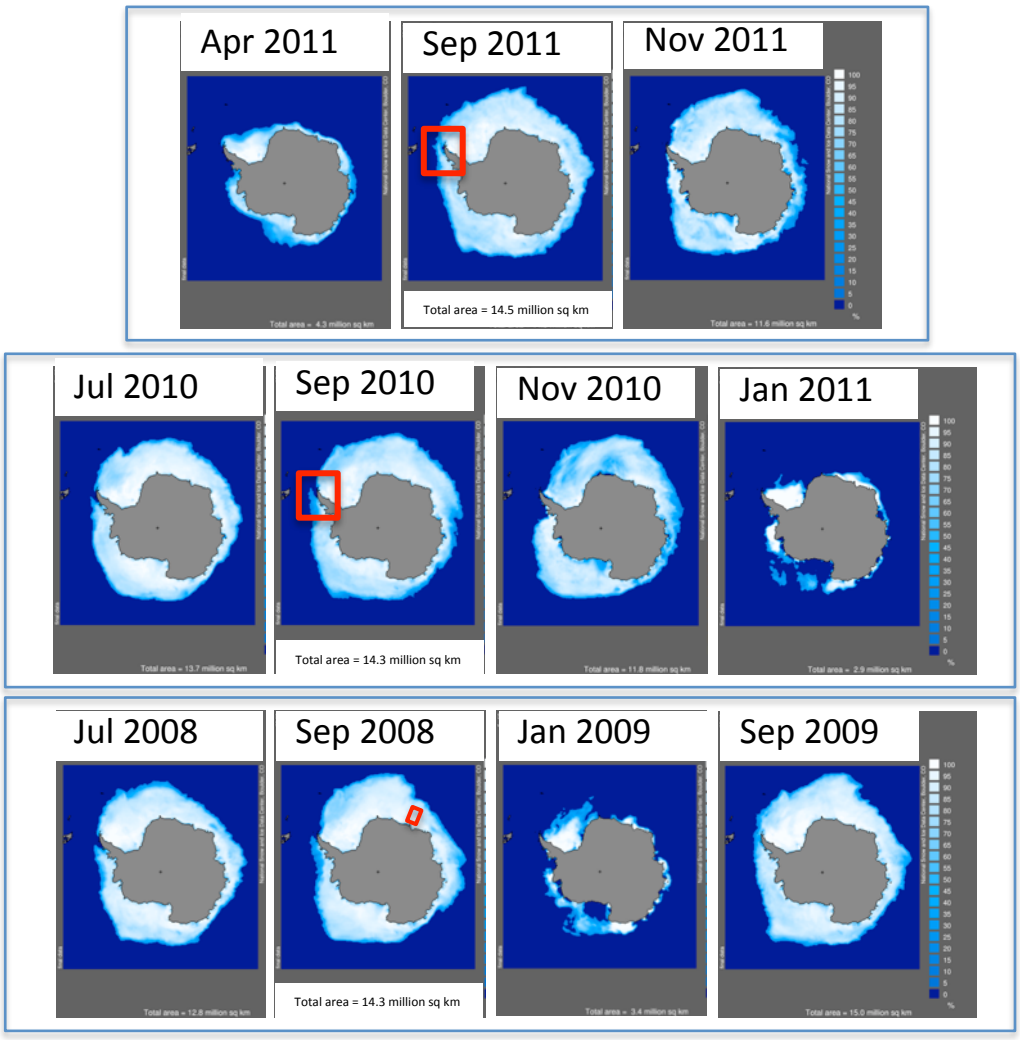


Figure 15. Monthly Antarctic sea ice concentration images

Chapter 3:

Transcriptome-wide profiles of gene expression of *Salpa thompsoni* in relation to variation of the pelagic environment of the Southern Ocean

ABSTRACT

The highly productive Western Antarctic Peninsula (WAP) region of the Southern Ocean supports dense populations of the Southern Ocean salp, *Salpa thompsoni.* This pivotal species in Antarctic marine ecosystems is subject to spatial and temporal heterogeneity as a consequence of variability in physical (temperature, salinity) and biological conditions (nutrients, chlorophyll, food availability) due to both seasonality and climate change. *Salpa thompsoni* shows marked variation in distribution, abundance, life stages, and timing of reproduction at seasonal to decadal scales, but the underlying drivers remain unclear. During this study, *S. thompsoni* was collected during research cruises in the WAP region (January and November 2011) and Indian Sector (February 2009). Transcriptome annotation for *S. thompsoni* matched sequences described for marine model species such as *Ciona intestinalis* (Tunicata), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata). Analysis of gene ontology showed good coverage of a wide range of gene ontology (GO) levels for each category for both the genome and transcriptome of *S. thompsoni.* Transcriptome-wide analysis of gene expression was analyzed with 2-way ANOVA for individuals collected across a range of Southern Ocean habitats (coastal, shelf, offshore), seasons, and years. All 2-way ANOVAs showed clear differentiation of salp samples from different regions, seasons, and samples. Differentially expressed genes were identified, and gene ontologies were examined to identify genes related to environmental stress responses. The identified stress-response genes showed different patterns of gene expression between seasons. Deep analysis of the *S. thompsoni* transcriptome revealed significantly different expression patterns – across 100s to 1000s of genes - from those in others areas (stations), regions, and seasons.

1. INTRODUCTION

Salpa thompsoni (along with krill) is one of the major grazers in the Southern Ocean. Salps are highly efficient, non-selective filter-feeders capable of ingesting particles between 1-1000μm (Madin et al., 1998). The repackaging of waste material in large, fast sinking fecal pellets means that *S. thompsoni* is a major source for vertical flux of organic matter (Dubischar et al., 1997; Huntley et al., 1989; Pakhomov, 2004; Pakhomov et al., 2004; Pakhomov et al., 2002; Perissinotto et al., 1998; Phillips et al., 2009).

Salpa thompsoni has been reported to occur in more frequent and more extensive blooms in some Southern Ocean regions in recent years (Loeb et al, 2012; Kokubun et al. 2013). In the last decades, a number of studies looking at historical data have shown a relationship between winter ice coverage, chlorophyll *a* and salp abundance (Atkinson et al., 2004; Pakhomov et al., 2006).

Salps belong to the Urochordata, one of the three subphyla of the Phylum Chordata. There have been studies showing impacts of environmental conditions on salps, such as high concentration of particulates in the water column (Pakhomov et al., 2003) and the link between winters with high sea ice coverage (Atkinson et al., 2004). However, few studies have investigated the influence of environmental conditions on salps in the context of genomic responses. Therefore, it is still unclear how individuals of this species respond physiologically and at a molecular level to environmental variation on both the seasonal and longer-term time-scales associated with climate change. High throughput sequencing has revolutionized whole-transcriptome analysis (i.e., patterns of gene expression based on RNA presence and quantity from a genome at any given time). The advances in next generation sequencing have made whole-genome and whole-transcriptome sequencing more accessible and available for non-model organisms. *Ciona*

intestinalis (Dehal et al., 2002) and *Oikopleura dioica* (Seo et al., 2001) are the closest relatives to *S. thompsoni* for which whole-genome sequencing has been completed and for which genome information is available. On the other hand whole-transcriptome (RNA-Seq) studies are also providing insights into how organisms are coping with different environmental conditions (Gallardo et al., 2014; Villarino et al., 2014; Smith et al., 2013) by analyzing patterns of gene expression under different conditions. Therefore these tools can be used to answer ecological questions and to understand the effects of different environmental conditions on the zooplankton communities experiencing seasonal or long-term variability in the environment.

This study reports the results of the application of whole-transcriptome sequencing analysis to provide new insights into the relationship between environmental variability and patterns of gene expression that may allow *S. thompsoni* to adapt to, and thrive in, changing environmental conditions in the Southern Ocean.

2. METHODS

2.1 Sample Collection

Salpa thompsoni was collected during three expeditions to two Southern Ocean regions in the Western Antarctic Peninsula (WAP) (austral spring and summer 2011) and the Indian Ocean (austral summer 2009). The austral spring 2011 cruise to the WAP region employed a Multiple Opening/Closing Net and Environmental Sensing System (MOCNESS; Wiebe et al., 1985) with a mouth opening of 1-m^2 and nine 335-µm mesh nets. Additionally, a 2.3 m^2 Isaacs-Kidd Midwater Trawl (IKMT; Isaacs and Kidd, 1953) with a 505 µm mesh net was employed (Cruise report LMG11-10; Bucklin et al., 2011). During the 2011 austral summer cruise (ANT-XVII/2) in the WAP region, the Rectangular Midwater Trawl (RMT 1+8) equipped with a real-time time-

depth-recorder (TDR) was used as standard gear to collect salp samples from the upper 200 m (Cruise report ANT-XVII/2; Fahrbach et al., 2011). In the Indian Sector during austral summer 2009, samples were collected also using a RMT 1+8. The RMT was towed obliquely in six different strata between the surface and 2,000 m.

For this study, and in order to compare across regions (WAP and Indian Sector), seasons (summer and spring), and areas (shelf, offshore and Bransfield Strait), four to five individuals were analyzed from each station selected; as a result a total of 48 individuals was used in this study (Table 1).

2.2 Handling of Salp Samples and Specimens

Salps species were identified under a dissecting scope and life stages were determined based on Daponte (2001), using the developmental stage of the oozoid embryo (when present) in the blastozooid or the stage of the development of the blastozooid chain in the oozoid. Each salp was measured using the standard oral-atrial length (Daponte et al., 2001) with a caliper. No abundance data are available, but the occurrence of *S. thompsoni* at the stations was recorded as presence / absence. For molecular analysis, the stomachs and embryos were removed by dissection to avoid contamination from prey; the remaining muscle tissue was flash-frozen in liquid nitrogen, stored at -80° C and shipped to the University of Connecticut for analysis.

2.3 RNA Extraction and Library Preparation

A total of 48 samples (individual zooids or organisms) was chosen from two different regions of the Southern Ocean (WAP and Indian sector) and three different seasons (Table 1) to represent a broad range of environmental conditions shown by PCA analysis (Fig. 2 and see Chapter 2). RNA was extracted from each whole animals using RNeasy system with on column DNase

treatment (Qiagen, Mansfield, MA). Resulting RNA was subjected to quality control on an Agilent 2100 Bioanalyzer, based on the RNA integrity number (RIN). The high quality RNA samples (RIN of 9.2 – 10) were used for construction of 2x100 paired end libraries with the TruSeq RNA Sample Prep, v2 protocol (Illumina, San Diego, CA), which uses poly-T oligo-attached magnetic beads to select for poly-A mRNA. To determine the functional concentration of each library (i.e., concentration of library fragments with adaptors), a qPCR assay was completed. Additionally the library was run on an Agilent 2100 Bioanalyzer to assess the size distribution of library fragments. Libraries were pooled based on the qPCR concentration to produce an equimolar concentration of functional fragments.

2.4 Sequencing

The 48 libraries were divided into 8 lanes with 6 samples per lane. Sequencing was performed according to the manufacturer's instructions on the Illumina HiSeq 2000 using v3.0 chemistry and the 2×100bp paired-end read mode (Illumina, San Diego, CA). The initial data analysis was conducted on the HiSeq 2000 instrument during the sequencing run. CASAVA 1.8.2 was employed to generate both report run statistics and the final FASTQ files comprising the sequence information. The latter was used for all subsequent bioinformatics analyses. Sequences were de-multiplexed according to the 6bp index code with 1 mismatch allowed.

2.5 Reference Transcriptome Assembly and Annotation

Trinity (v. r20140413p1) was used to build the reference transcriptome based on 40 samples, the eight remaining samples were not used because there was a chance they had an embryo, which would add background noise to the assembly. Approximately 500M total reads were used (~12.5M/sample). From the assembly step, 217,849 total contigs were screened for bacterial contamination using Blastn against all bacterial sequences

(ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/all.fna.tar.gz) and an Evalue ≤ 0.001.Transcripts in the assembly were annotated using BLASTX against the NR protein database with an Evalue ≤ 0.001. Blast2Go was used to identify putative gene function based on homology to known transcript sequences (Conesa et al., 2005).

2.6 Differential Gene Expression

Fastq expression data were analyzed using BWA pipeline via the GeneSifter analysis software (http://genesifter.net, Geospiza, Inc., Seattle, USA) comparing the unique samples with the new transcript reference to generate expression values. Reads Per Kilobase per Million (RPKM) values were used to normalize the data by calculating the number of reads mapped to each contig divided by the number of million reads mapped, then divided by the total length of contigs in kilobases (Kb).

2.7 Statistical Analysis

2.7.1. Bray-Curtis similarity

To analyze similarity of gene expression among the 48 samples, a matrix of RPKM values from each 216,931 transcripts for each sample was used to calculate similarity with the Bray-Curtis function in the FATHOM toolbox (Jones, 2014) in MATLAB v. R2013a (The MathWorks, Inc., Natick, MA, USA).

2.6.2. Two-way ANOVA analysis

Comparisons of gene expression levels (i.e., transcriptome RPM data) among regions (WAP vs Indian Sector), seasons (austral summer vs austral spring) and areas (shelf vs offshore) for the

were done using a parametric two-way Analysis of Variance (ANOVA) tests, assuming equal variance, and designated False Discovery Rate (FDR) values (Benjamini et al., 1995) in GeneSifter (Geospiza, WA). The two-way ANOVAs were done using 4x (commonly used in the literature), 20x and 100x (highly significant differential expression) as the cutoff for fold-changes in up-regulated and down-regulated genes. The result from each analysis was exported as a list of genes that were differentially expressed under different conditions.

2.7.3 Enrichment analysis

Enrichment analysis was performed comparing the reference transcriptome and the list of genes resulting from each two-way ANOVA analysis (with a cutoff of 4x). Gene Ontolgy (GO) term enrichments revealed groups of interesting genes when compared with a reference group. This functionality was introduced in Blast2GO by integrating Gossip (Blüthgen *et al*., 2004). Gossip computes Fisher's Exact Test applying robust FDR (false discovery rate) correction for multiple testing and returns a list of significant GO terms ranked by their corrected or one-test *P*-values (Conesa et al., 2005). In this study, the enrichment analysis was used to find the GO terms associated with the genes that are over-represented using the annotations from the reference transcriptome.

3. RESULTS

3.1 Sequencing and Assembly

Hi-Seq Illumina sequencing resulted in 2.32 - 11.21GB (i.e., 10^9 base pairs) per sample, or a total of 23.16 - 112.07 M reads, of which 91% had quality (Q) values > 30 (i.e.,the chances that a

given base is called incorrectly is 1 in 1000). After adaptor removal and quality control, the Trinity/BLAST pipeline produced 217,849 transcripts with an average length of 709.02 bp. After the bacterial screening, 918 contigs that mapped to the bacterial reference with an E-value of <0.001 were removed to avoid any misinterpretation of the data. The remaining 216,931 contigs were used for downstream analysis.

3.2 Reference Transcriptome Annotation

From the reference transcriptome assembly, 217,849 transcripts were submitted for annotation based on similarity to genes within the NR protein database using an Evalue $\leq$ 0.001. A total of 41,347 contigs (18%) was successfully annotated

3.3 Gene Ontology

Out of the 216,931 transcripts from the assembly, only 41,210 sequences associated with predicted, hypothetical and known proteins were submitted to Blast2GO for further analysis. A total of 14,377 sequences had blast hits; 13,058 were successfully mapped and 13,775 were annotated (Fig. 3). The transcripts had blast hits matching marine species in the database, the top-hit species were *Ciona intestinalis* (Tunicate), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata) (Fig. 5). Gene ontology analysis of the reference transcriptome showed a wide distribution of GO terms across all levels (level 2-13) for the three categories, biological process (BP), molecular function (MF) and cellular component (CC) (Fig. 4).

3.4 Differential Expression

3.4.1 Bray-Curtis similarity

The 48 samples had a high degree of variability; Bray-Curtis similarity analysis showed 55-95 % similarity (Fig. 6), with some patterns that were regionally or temporally specific. Highest similarities were found between samples collected in the WAP in the spring of 2011. This group of samples was more similar to samples collected in the Indian sector during summer 2009 than those from the WAP during summer 2011 (Fig. 6). It is important to mention that there were differences in life stage frequencies among seasons; all the samples collected during the summer were individuals in stage 2 of their life cycle, while the samples collected during the spring corresponded to the earlier life stage 0. The difference in life stage of the individuals can contribute to the differences in gene expression patterns between seasons.

3.4.2 Two-way ANOVA analysis.

Two-way ANOVA analysis showed significant differentiation among areas within a region and between seasons. Figure 7 shows the clear differentiation between the two areas of the WAP (Bransfield Straight and an offshore station) where salps were collected during the austral spring 2011. This holds true regardless of the gene expression fold change (4x, 20x or 100x) cutoff used. The offshore station (Stn. 4) showed the most differentiation, followed by Stn. 12, which was the furthest offshore station in the Bransfield Straight. Gene expression patterns for Stns. 14 and 15 were similar.

A similar analysis was done to explore larger scale regional differences by comparing gene expression of salps collected from the WAP and Indian sector during the austral summer. There were distinct and clear differences between the regions (Fig. 8). The offshore Stn. 184 showed a clear differentiation with the other offshore station (Stn. 165), while the shelf stations showed similar patterns of differentially expressed genes. It is important to mention that salps used in

these two-way ANOVAs were at the same life stage, therefore reducing the potential for differences in gene expression patterns due to differences in their life stages.

When the two-way ANOVA was used to compare gene expression of salps collected in one region (WAP) in different seasons (austral spring and summer), there was clear differentiation in the gene expression patterns between the austral spring and summer (Fig. 9). In contrast, there was little differentiation among salps collected during the austral summer (Fig. 9), also evident from the similarity Bray-Curtis heat map (Fig. 6). Samples collected at Stn. 4 during the austral spring again showed differentiation when compared to the rest of the samples collected during the same season in the Bransfield Strait (Fig. 9). For this analysis *S. thompsoni* life stages were different between the summer and spring and therefore could be a source for the different patterns of gene expression.

Finally, two-way ANOVA comparison between samples collected from the WAP (summer 2011) and Indian Sector (summer 2009) was made (Fig. 10). Gene expression patterns were similar within the shelf and offshore WAP samples, while the Indian Sector samples showed differentiation between them when a 20x fold change was used. However, these patterns were all absent when a 100x fold change threshold was used for the two-way ANOVA analysis.

3.4.3. Enrichment analysis

Enrichment analysis in Blast2GO showed a change in the percentage of GO terms, meaning that – when compared to the reference transcriptome – differential expression under different environmental conditions had an effect on the proportion of GO terms been expressed. All the GO categories [cellular component (CP), molecular function (MF) and biological process (BP)]

were present in the over-represented GO terms. This analysis showed that only a small fraction of genes in each GO category were differentially expressed (Table 3).

Differentially expressed genes during the austral spring of 2011 in the WAP region were associated with six GO terms, primarily within MF (70%), while the rest of the GO terms were associated with BP (30%) (Fig. 11). In contrast, during austral summer in the WAP region 2011, salps had 57 GO terms associated with the differentially expressed genes, of which 99% were associated with BP, while only 1% was associated with MF (Fig. 12). Finally, the differentially expressed genes from a two way ANOVA comparing austral summer and austral spring in the WAP included 41 GO terms, which were associated with BP (64%), MF (8%) and CC (28%) (Fig. 13).

4. DISCUSSION

The number of samples and the sequencing effort produced a complete reference transcriptome. The statistical analysis of transcriptome-wide patterns of gene expression revealed differences among salps collected in different environmental conditions.

Gene ontology analysis showed a good coverage of a wide range of GO levels for each of the categories for both the transcriptome and genome sequencing efforts for *S. thompsoni.* Even though the transcriptome data set is more robust – and therefore will offer more coverage – it is noteworthy that the genome data shows similar distribution across the different categories and levels of gene ontology. Similarly the top blast hits for the transcriptome annotation matched sequences described for marine model species such as *Ciona intestinalis* (Tunicata), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata). This clearly shows that the new genomic and transcriptomic catalog of genes provided here is a valid

basis for new 'omics efforts to understand how marine organisms are reacting to different environmental conditions at a molecular level.

When samples collected in different regions, seasons or areas were compared using similarity analysis of individual RPKM values per transcript, there were high levels of similarity among samples collected during the austral spring in the WAP. However the offshore station (Stn. 4) showed lower similarity to the other samples from the WAP austral spring 2011 collections. On the other hand, RPKM values for samples that were collected during the austral summer showed lower similarity, and the WAP region showed the least similarity of the three cruises. This might be due to sampling different life stages, since the samples that were collected in WAP and Indian Sector during the austral summer belong to the same life stage (life stage 2) (See Table 1), whereas samples collected during the austral spring in the WAP belong to and earlier life stage (life stage 0).

The two-way ANOVAs showed clear differentiation of gene expression patterns of samples collected in different areas of the WAP during the austral spring of 2011. This differentiation held true for gene expression thresholds of 4x and 20x. However, when the gene expression threshold was raised to 100x, the differentiation between Stns. 14 and 15 in the Bransfield Strait was not as clear. The gene expression patterns for the offshore station (Stn. 4) were distinguished from the rest of the samples collected in the WAP during austral spring 2011 regardless of the fold-charge threshold used. A different gene expression pattern was observed in the two-way ANOVA comparing samples collected during the austral summer from the WAP. At a low fold-change threshold (4x), samples collected on the shelf were not differentiated, while the majority of the samples collected in offshore areas (Stns. 184 and 165) clustered together. When the threshold was increased to 20x or 100x, all except for one of the shelf stations clustered together.

The same was true for the offshore stations. When comparing samples collected in the WAP region during the austral spring and summer 2011, two clusters were formed: one for each collection event. Samples collected in austral summer did not form individual clusters, while the samples from austral spring did at both 20x and 100x fold-change thresholds. This clear differentiation between summer and spring could be due to the different life stages of the salps collected during the different seasons. Finally a two-way ANOVA comparing samples of the same life stage, collected during the austral summer in WAP and Indian Sector, showed a clear differentiation between samples from the two regions at 20x fold-change threshold. At a higher threshold (100x), samples from Stn. 35 in the Indian Sector and offshore Stn. 165 remained clustered individually, while a sample from the WAP shelf (Stn. 161) and Stn. 33 in the Indian Sector clustered together. All of the two-way ANOVAs showed clear differentiation – albeit in different patterns – of the salps samples from different regions, seasons, and samples. This finding indicates that samples show similar gene expression patterns among samples in the area (station), with significant differentiation of transcriptome-wide gene expression patterns among samples collected during different seasons and in different regions.

The enrichment analysis for the genes differentially expressed during the austral spring in the WAP had few GO terms associated with them; most were related to tyrosine (an amino acid used to synthesize proteins). This could be due to the earlier stages of the WAP population and could be related to the emergence of *S. thompsoni* from overwintering. During the summer, the genes differentially expressed had a wider range of GO terms, most of them associated with the GO category of BP (i.e., regulation of response to stress, regulation of programed cell death, cellular response to stress). This can be explained by the fact that these samples were collected at the end of January, during a period of rapid environmental change in the seasonal shift from summer to

fall. Also, the salp population had a wide range of life stages (see Chap. 2), so it is possible that the more developed life stages have a wider range of expressed genes. When a similar analysis was done with the genes differentially expressed for both summer and spring 2011, there was also a wide range of GO terms, however this time all three categories were represented: BP (64%), MF (8%), and CC (28%). Finally stress-response genes were identified and showed different patterns of gene expression between seasons however next steps should involved genes related to reproduction and feeding to further understand the response of *S. thompsoni* to different environmental conditions.

This study is the first transcriptome sequencing effort for this group of tunicates (salps). The resulting 217,000 transcripts provide a valuable tool for biological and bioinformatics inferences. In conjunction with the reference genome (Chap. 1), these data will enable and facilitate novel analytical approaches to understanding how an important Antarctic species may be affected by environmental variation. Due to the increase of *S. thompsoni* in the last decades in high latitudes of the Southern Ocean, it is critically important to understand and seek to predict how the species may respond to the changing environmental conditions, including warming and ocean acidification. Our findings revealed differing patterns of gene expression associated with different seasons, , life stages and regions. Future research should involve looking at responses across the pelagic community, including krill and other key species, to fully understand the present and future effects of the changing Antarctic environment on the ecosystem.

Figure Legends:

Figure 1. Map of sampling domains and cruise tracks. (A) WAP region in austral spring (2011); (B) WAP region in the austral summer (2011); and (C) Indian Sector in austral summer (2009). Offshore station (green) and Shelf or Bransfield stations (red) where samples were chosen for the RNA-Seq analysis.

Figure 2. Principal component analysis (PCA) of environmental data from stations surveyed in the WAP region during the austral spring 2011 (red) and the austral summer 2011 (blue); and in the Indian sector, surveyed during the Austral summer 2009 (green). Salp presence (solid) or absence (x) is also shown for each collection.

Figure 3. *Salpa thompsoni* transcriptome and genome data distribution in Blast2GO

Figure 4. GO-Level distribution of *S. thompsoni* transcriptome and genome data

Figure 5. Top-Hit species data distribution in Blast2GO

Figure 6. Heat map showing Bray-Curtis similarity of the gene expression patterns among 47 samples.

Figure 7. Results of two-way ANOVA comparing regions, stations and individuals during the austral spring in the WAP. The threshold fold-change for differential gene expression was set at 4x (left), 20x (center) and 100x (right). A blue frame surrounds offshore stations while red lines frame shelf or Bransfield Strait stations. Gene expression patterns with up-regulated (green) and down-regulated (red) are reflected in the clustering.

Figure 8. Sample and gene clustering from results of a two-way ANOVA comparing, comparing regions, stations and samples for the austral summer in the WAP. The threshold fold-change for

differential gene expression was set at 4x (left), 20x (center) and 100x (right). A blue frame surrounds offshore stations while red lines frame shelf or Bransfield Strait stations.

Figure 9. Sample and gene clustering based on results of a two-way ANOVA for the austral sprin 2009 (Indian Sector) and austral summer 2011 (WAP); comparing regions, stations and individuals. The threshold fold-change for differential gene expression was set at 20x (left) and 100x (right). A blue frame surrounds offshore stations while red lines frame shelf or Bransfield Strait stations.

Figure 10. Plot showing sample and gene clustering from results of a two-way ANOVA for the austral summer and austral spring of 2011 in the WAP region; comparing regions, stations and individuals. The threshold fold-change for differential gene expression was set at 20x (left) and 100x (right).

Figure 11. Over-represented GO terms in *S. thompsoni* collected during the austral spring in the WAP region

Figure 12. Over-represented GO terms in *S. thompsoni* collected during the austral summer in the WAP region

Figure 13. Over-represented GO terms in the WAP (austral spring vs austral summer) compared to the GO terms from the reference transcriptome

Tables:

Table 1. Experimental design including two regions (WAP and Indian Sector), three seasons, two areas, and four to five individuals sampled at each station.

Region	Season	Area	Station	Life stage 0	Life stage II	Number of salps (replicates)
Indian Sector	Summer 2009	Offshore	33	0	4	4
Indian Sector	Summer 2009	Offshore	35	0	4	4
WAP	Summer 2011	Shelf	161	0	5	5
WAP	Summer 2011	Shelf	174	0	5	5
WAP	Summer 2011	Offshore	184	0	5	5
WAP	Summer 2011	Offshore	165	0	5	5
WAP	Spring 2011	Bransfield Strait	14	5	0	5
WAP	Spring 2011	Bransfield Strait	15	5	0	5
WAP	Spring 2011	Bransfield Strait	12	5	0	5
WAP	Spring 2011	Offshore	4	5	0	5

Table 2. Two-way ANOVA analysis of four different comparisons using the Benjamini-Hochberg correction.

	Fold change		
Two-way ANOVA	**4**	**20**	**100**
Spring 2011 - WAP	7071	833	83
Summer 2011 - WAP	5407	173	8
WAP 2011 (summer vs spring)	24914	1821	179
Summer 2009 – 2011 (Indian Sector vs WAP)	16315	536	25

Table 3. Number of genes showing significant enrichment based on Fisher's exact test of differentially expressed genes for three different comparisons in the WAP region.

Enrichment Analysis		Biological Process		Molecular Function		Cellular Components	
		Genes	GO terms	Genes	GO terms	Genes	GO terms
WAP	Test	1582	27	203	6	569	8
	Reference	29975		1810		9537	
Spring - 2011 WAP	Test	11	3	26	3	0	0
	Reference	35		96		0	
Summer – 2011 WAP	Test	218	56	3	1	0	0
	Reference	3483		127		0	

Region	Season	Area	Station	Life stage 0	Life stage II	Number of salps (replicates)
Indian Sector	Summer 2009	Offshore	33	0	4	4
Indian Sector	Summer 2009	Offshore	35	0	4	4
WAP	Summer 2011	Shelf	161	0	5	5
WAP	Summer 2011	Shelf	174	0	5	5
WAP	Summer 2011	Offshore	184	0	5	5
WAP	Summer 2011	Offshore	165	0	5	5
WAP	Spring 2011	Bransfield Strait	14	5	0	5
WAP	Spring 2011	Bransfield Strait	15	5	0	5
WAP	Spring 2011	Bransfield Strait	12	5	0	5
WAP	Spring 2011	Offshore	4	5	0	5

Table 1. Experimental design including two regions (WAP and Indian Sector), three seasons, two areas, and four-to-five individuals sampled at each station.

Two-way ANOVA	Fold change 4	20	100
Spring 2011 - WAP	7071	833	83
Summer 2011 - WAP	5407	173	8
WAP 2011 (summer vs spring)	24914	1821	179
Summer 2009 – 2011 (Indian Sector vs WAP)	16315	536	25

Table 2. Two-way ANOVA's for four different comparisons. Benjamini-Hochberg correction.

Enrichment Analysis		Biological Process		Molecular Function		Cellular Components	
		Genes	GO terms	Genes	GO terms	Genes	GO terms
WAP	Test	1582	27	203	6	569	8
	Reference	29975		1810		9537	
Spring - 2011 WAP	Test	11	3	26	3	0	0
	Reference	35		96		0	
Summer – 2011 WAP	Test	218	56	3	1	0	0
	Reference	3483		127		0	

Table 3. Enrichment analysis (Fisher's exact test) of differentially expressed genes for three different comparisons in the WAP region.

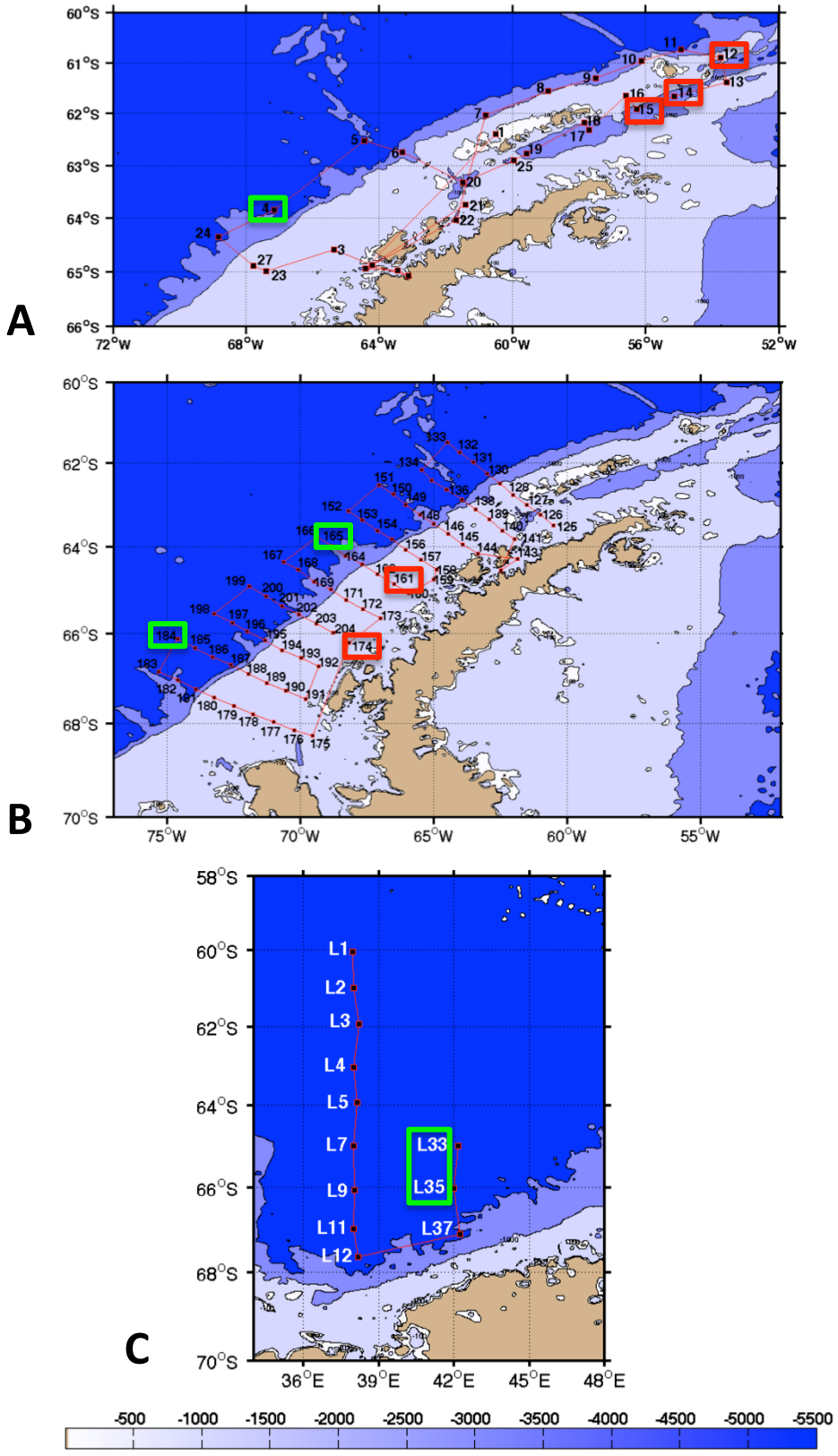


Figure 1. Map of sampling domains and cruise tracks.

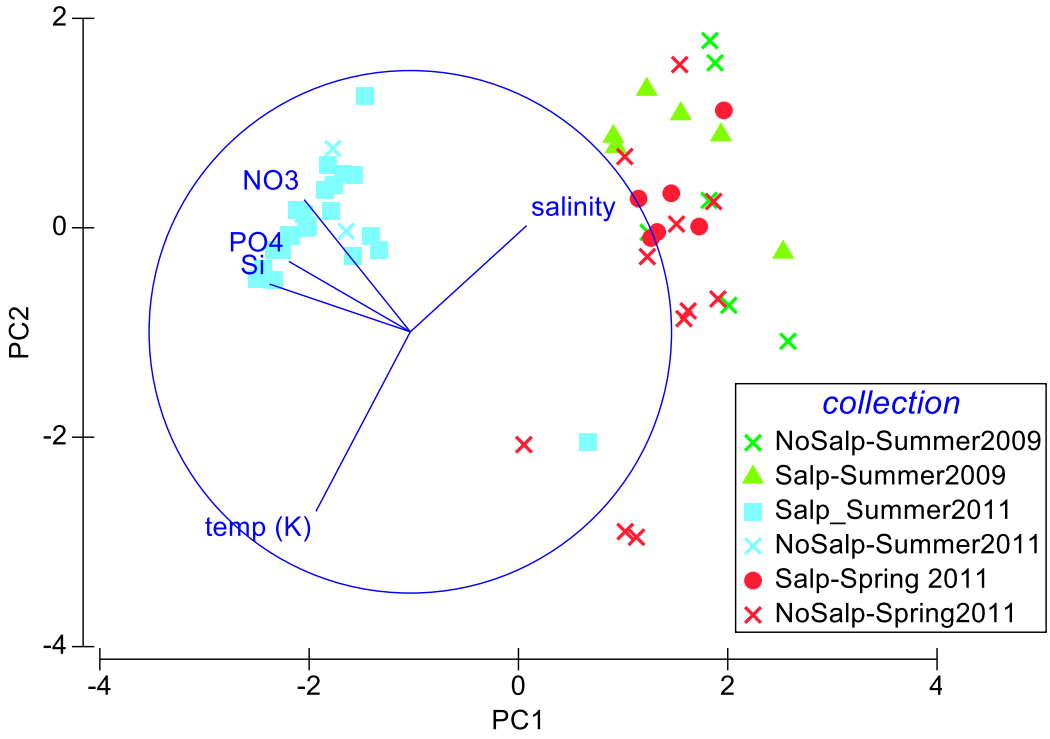


Figure 2. – Principal component analysis (PCA) of environmental data

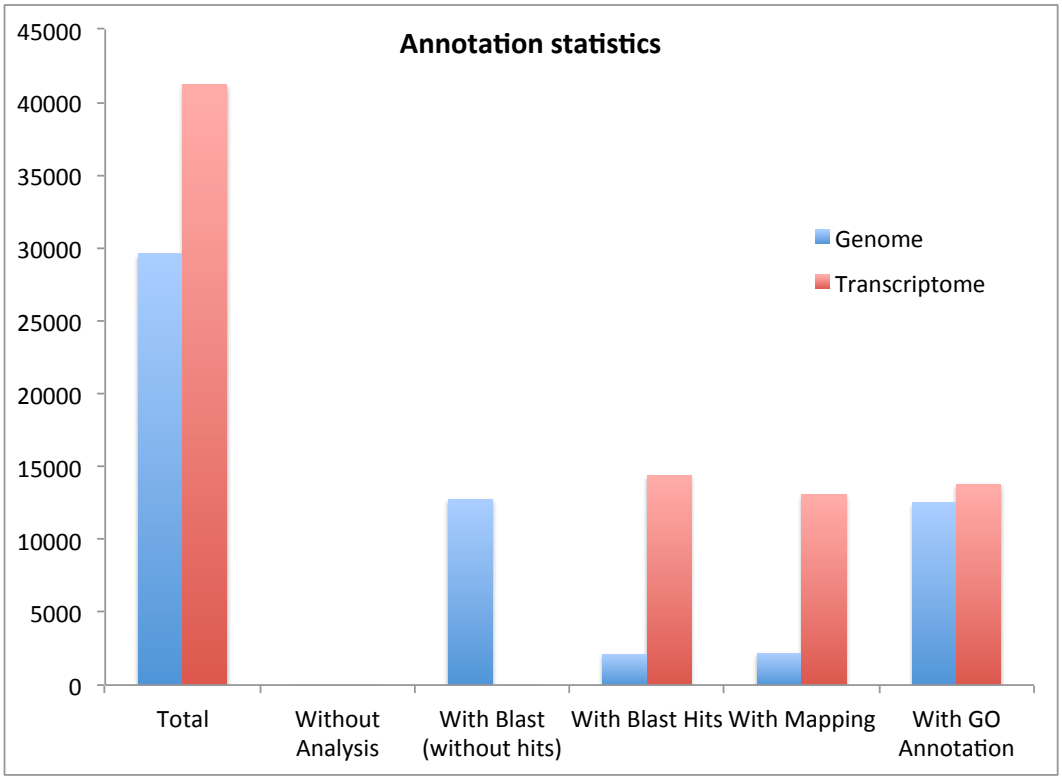


Figure 3. *Salpa thompsoni* transcriptome and genome data distribution in Blast2GO

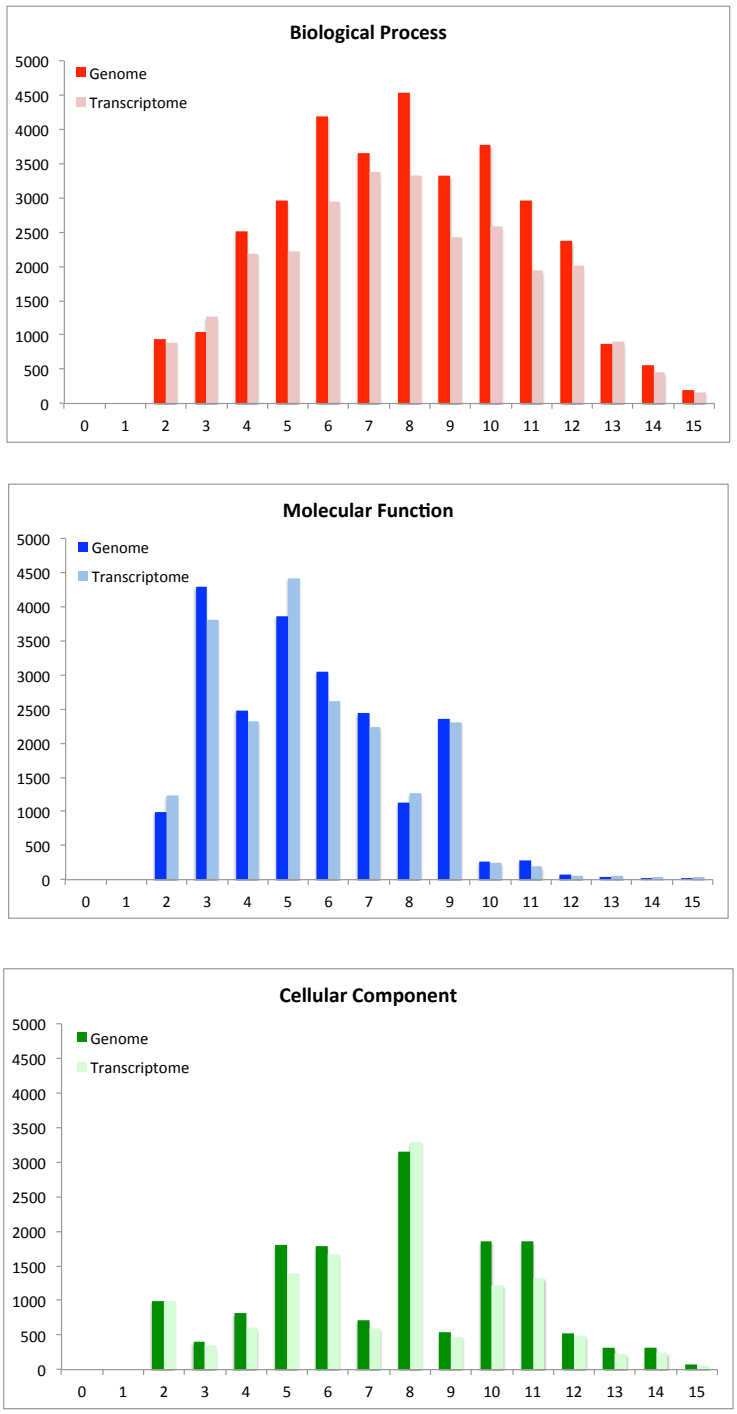


Figure 4. GO-Level distribution of *S. thompsoni* transcriptome and genome data

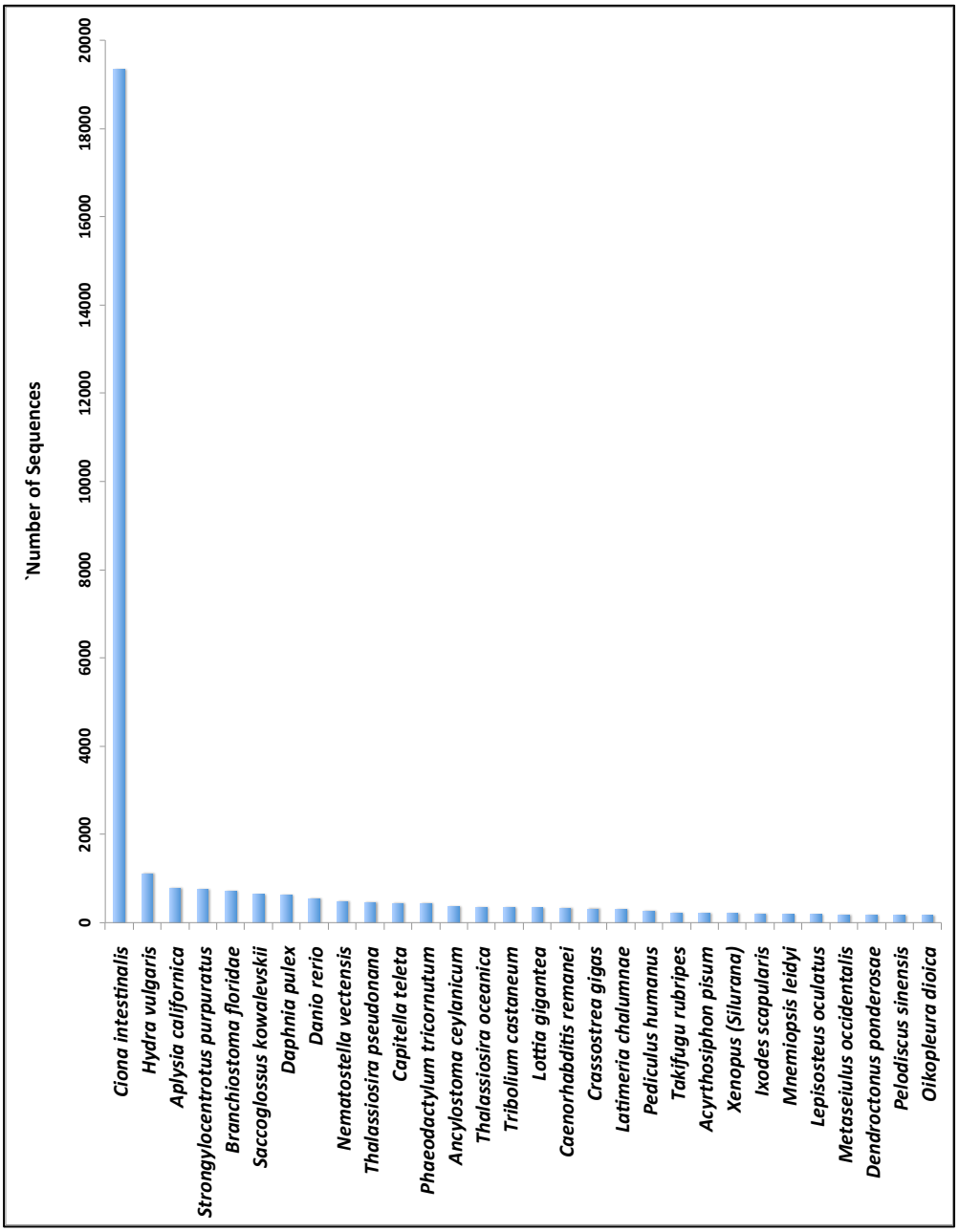


Figure 5. Top-Hit species data distribution in Blast2GO

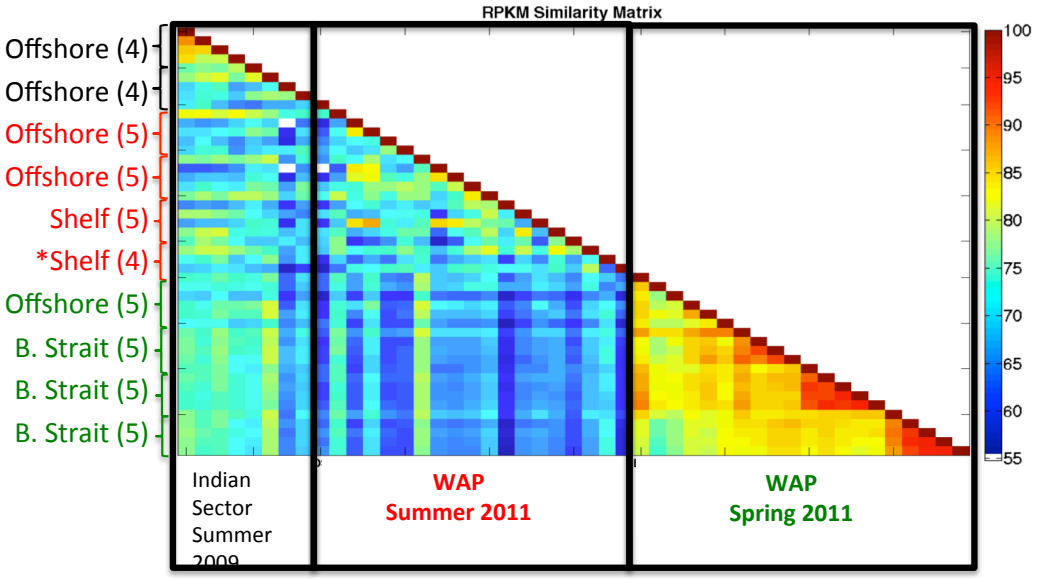


Figure 6. Heat map of Bray-Curtis similarity of RPKM values.

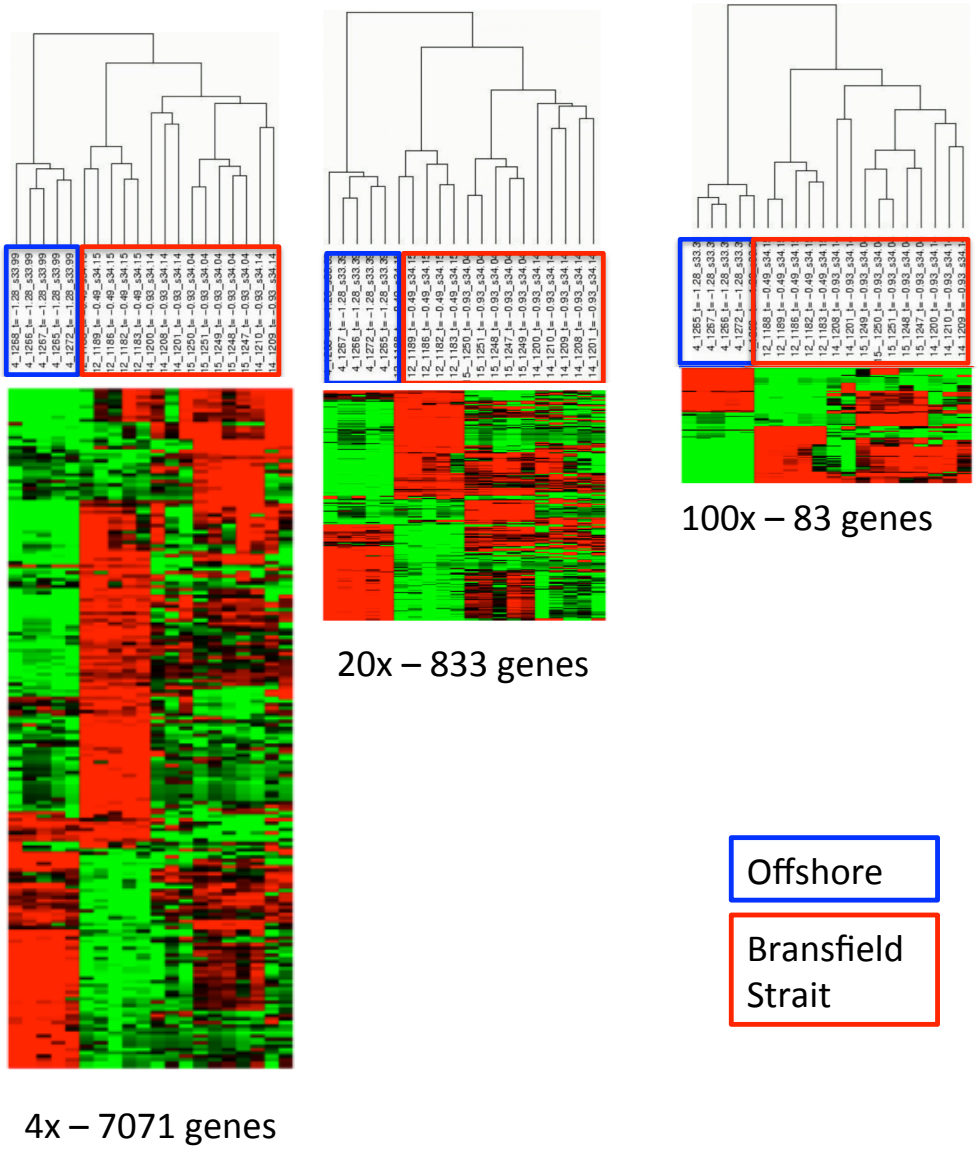


Figure 7. Sample and gene clustering from results of a Two-way ANOVA of samples collected during the austral spring in the WAP region.

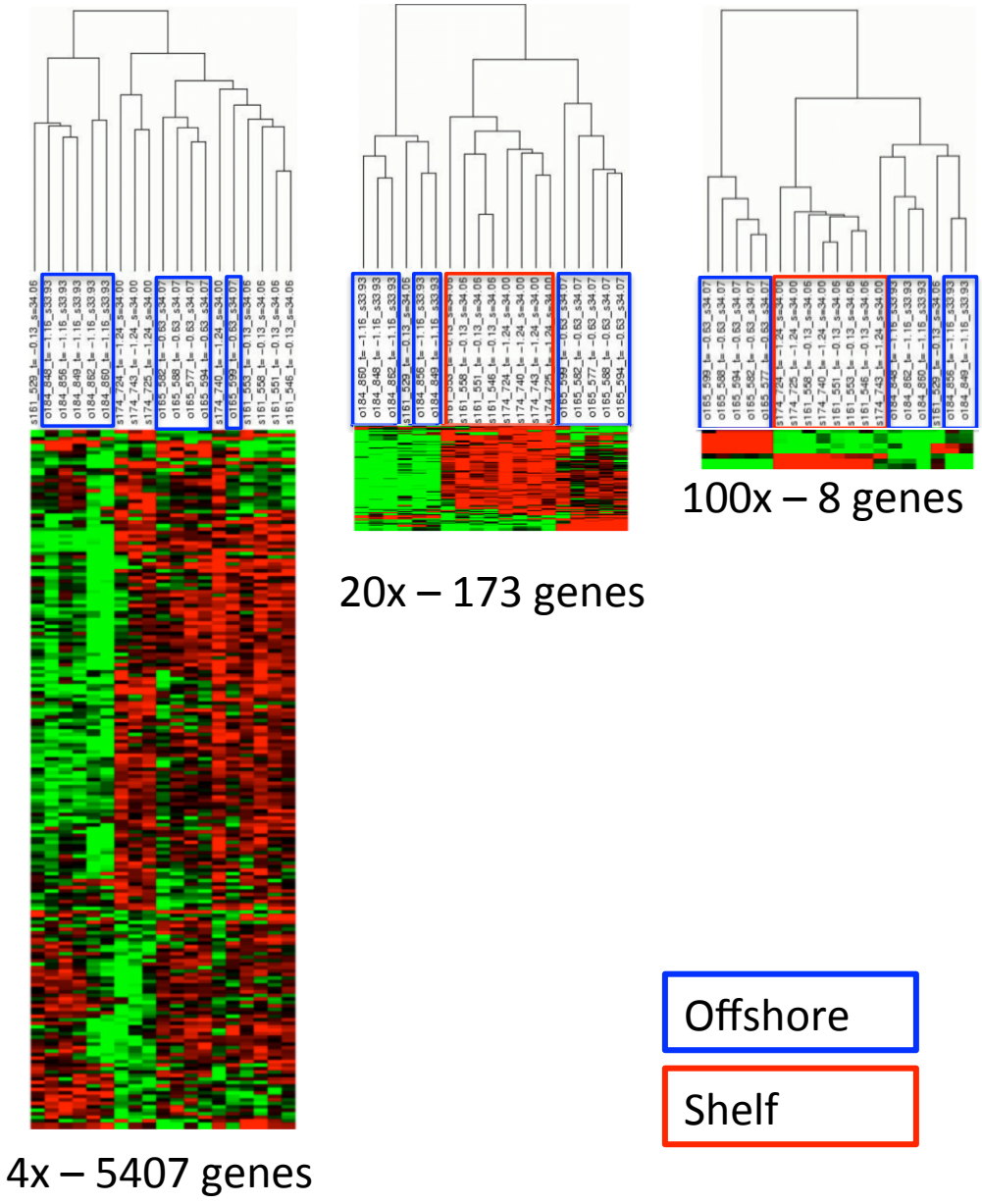


Figure 8. Sample and gene clustering from results of a Two-way ANOVA of samples collected during the austral summer in the WAP region.

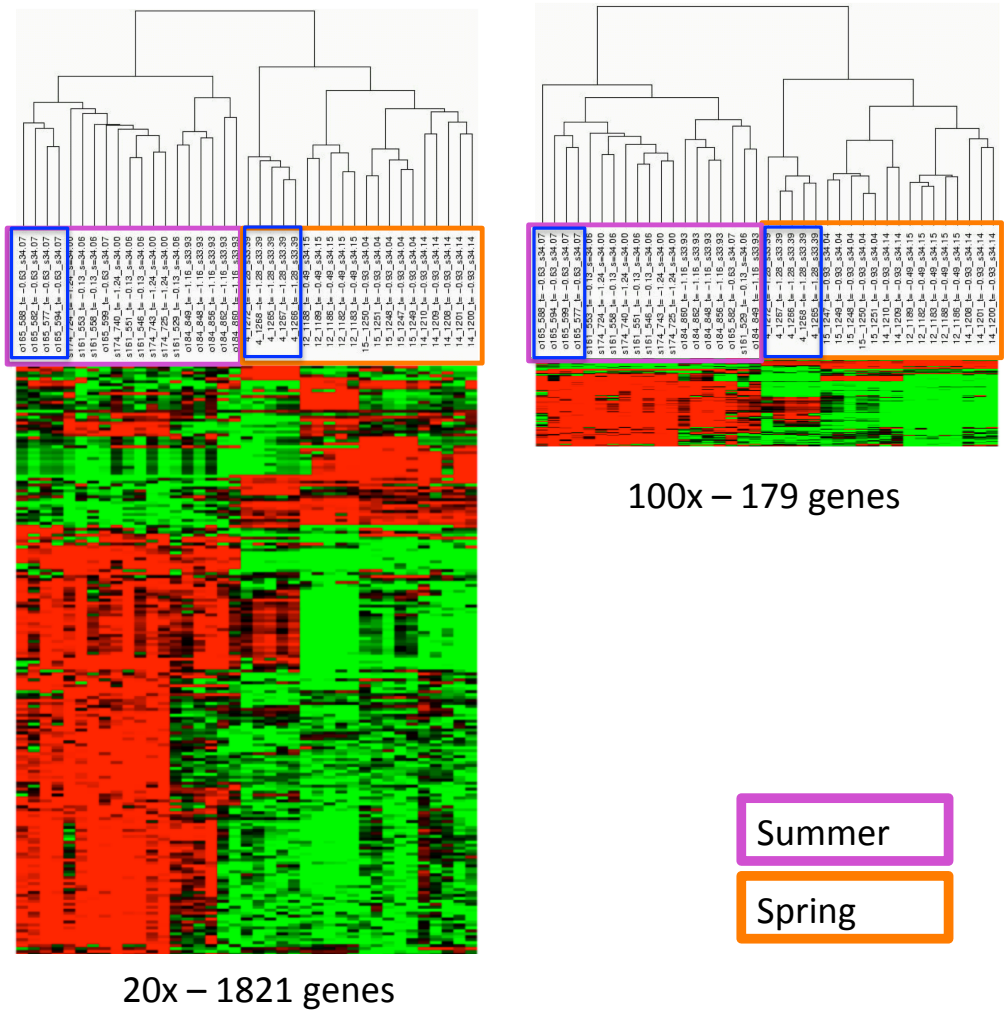


Figure 9. Plot showing sample and gene clustering from results of a two-way ANOVA for the austral summer and austral spring of 2011 in the WAP region .

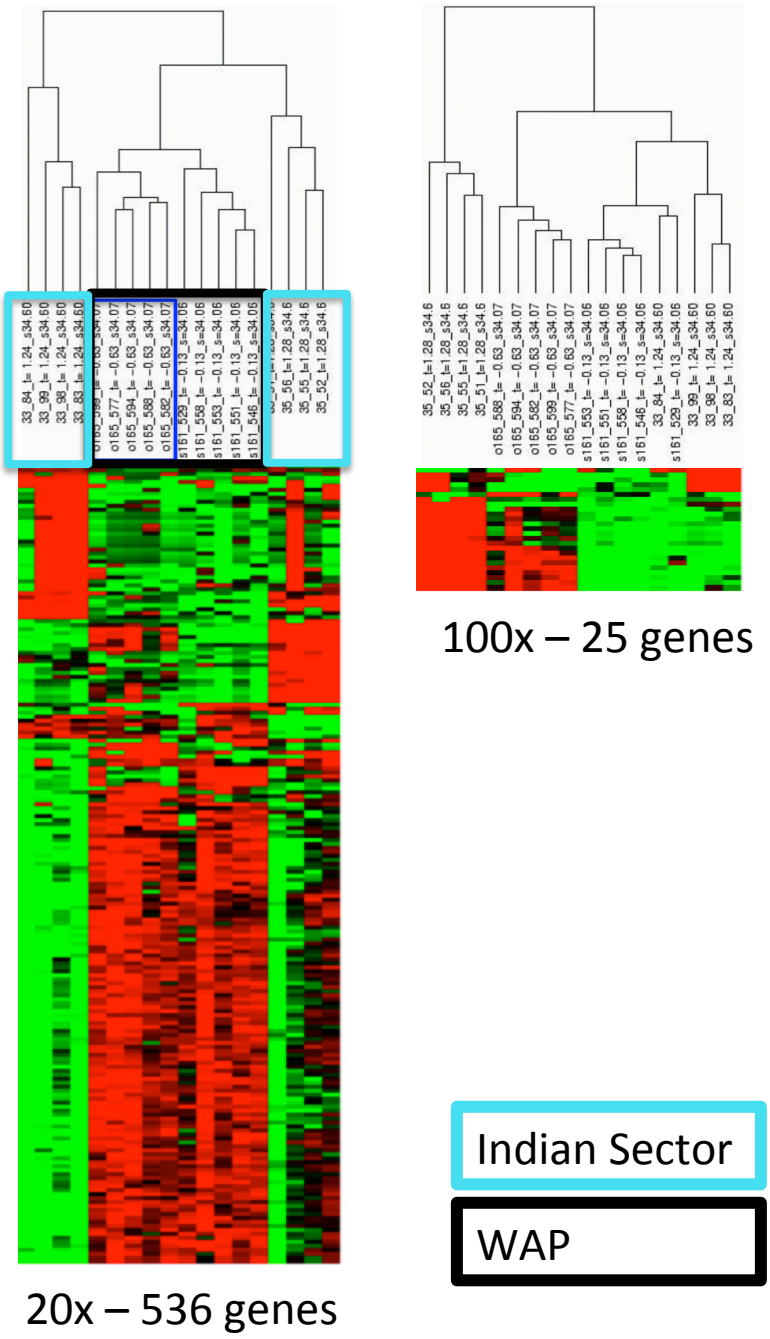


Figure 10. Plot showing sample and gene clustering for the austral summer 2009 (Indian Sector) and austral summer 2011 (WAP).

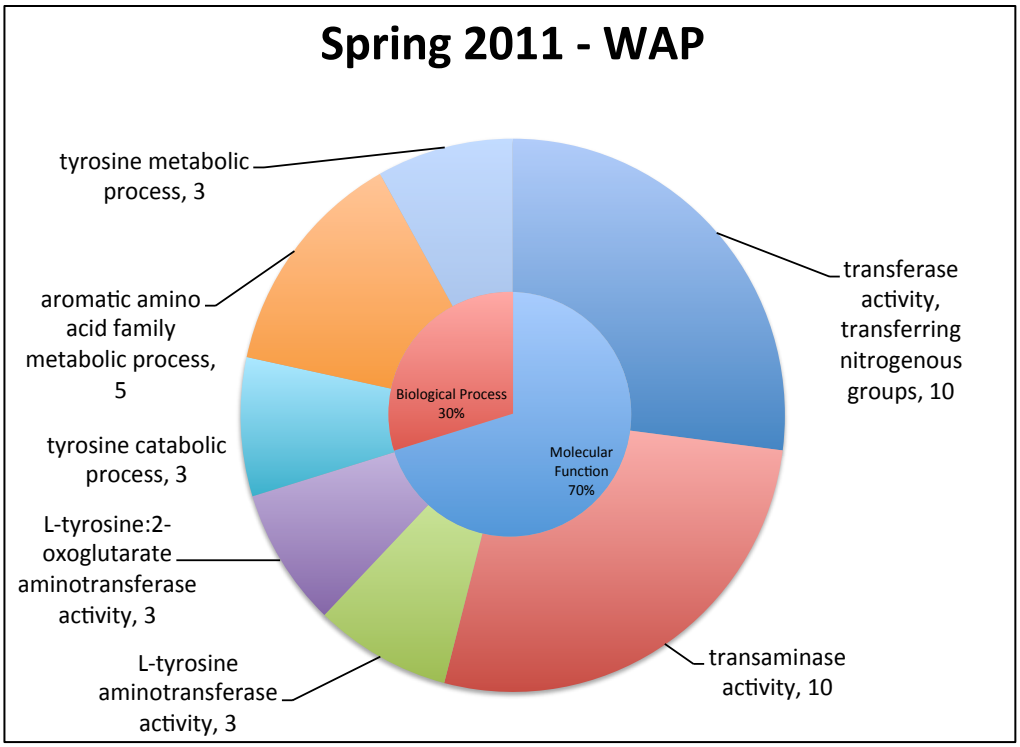


Figure 11. Over-represented GO terms in *S. thompsoni* collected during the austral spring in the WAP region

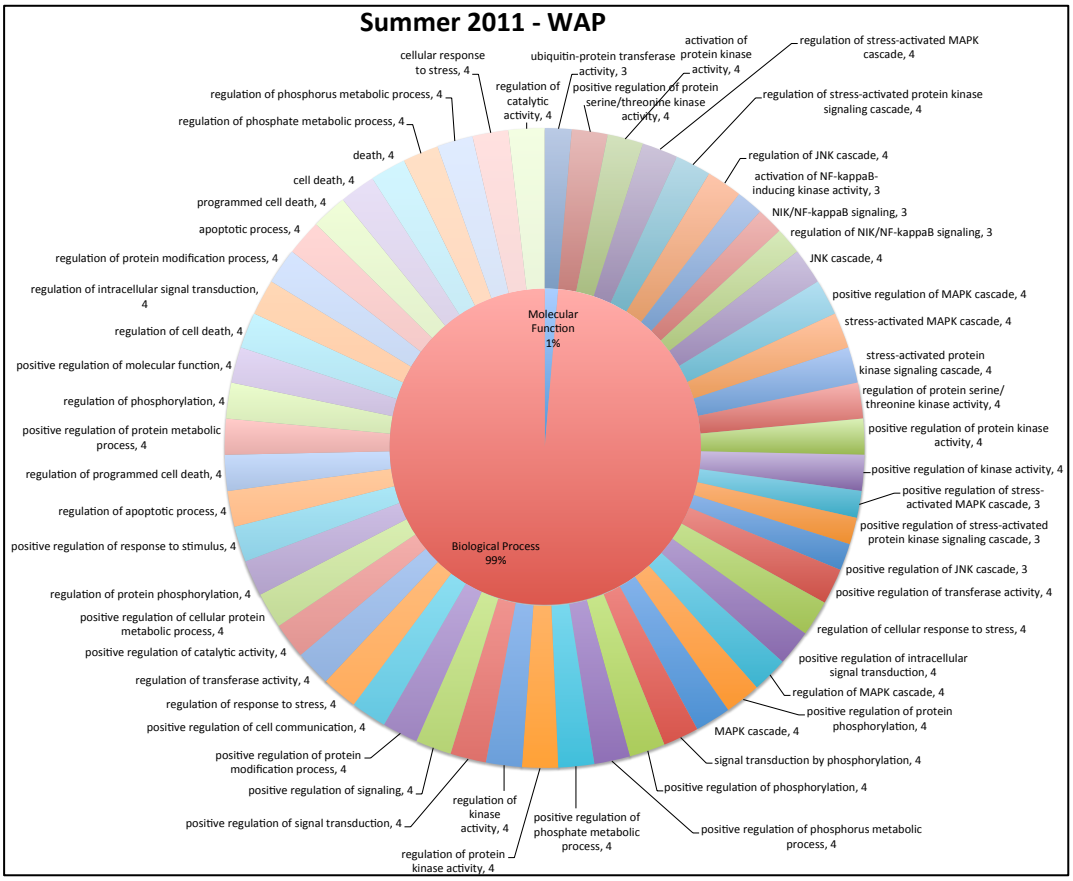


Figure 12. Over-represented GO terms in *S. thompsoni* collected during the austral summer in the WAP region

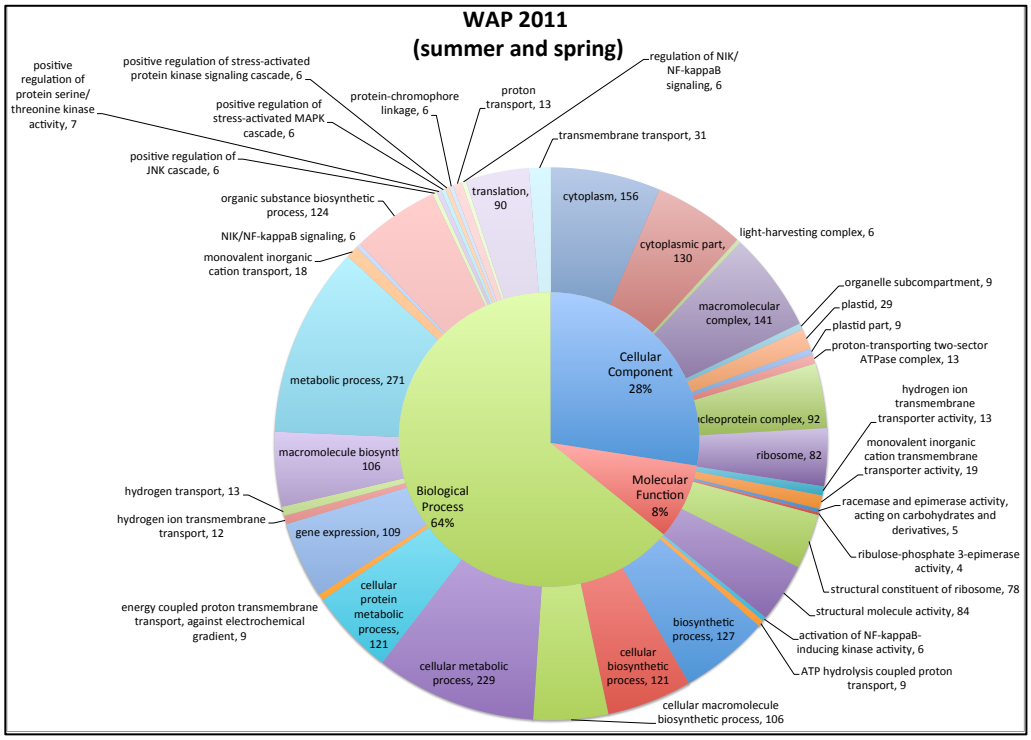


Figure 13. Over-represented GO terms in the WAP (austral spring vs austral summer)

Chapter 4:

Population genetic variation of the Southern Ocean krill, *Euphausia superba*, in the Western Antarctic Peninsula region based on mitochondrial single nucleotide polymorphisms (SNPs)

(This section is a manuscript published in Jounral Deep Sea Research Part II: Topical Studies in Oceanography. Batta-Lona, P.G., Bucklin, A., Wiebe, P.H., Patarnello, T., and Copley, N.J. (2011) Population genetic variation of the Southern Ocean krill, Euphausia superba, in the Western Antarctic Peninsula region based on mitochondrial single nucleotide polymorphisms (SNPs). *Deep Sea Research Part II: Topical Studies in Oceanography* 58, 1652-1661)

ABSTRACT

Population genetic diversity and structure of the Antarctic krill *Euphausia superba* was examined for collections from the Western Antarctic Peninsula (WAP) region made during US GLOBEC cruises in austral Fall, 2001 and 2002. The molecular markers used were allele frequencies at single nucleotide polymorphism (SNP) sites in the gene encoding mitochondrial cytochrome B (CytB). Four 4-fold degenerate SNP sites were selected, and allelic variation was detected using a multiplexed single-base extension PCR protocol for a total of 585 *E. superba*, including all 6 furcilia larval stages, juveniles, and adults). Confirmation of species identification was done using a competitive multiplexed species-specific PCR (SS-PCR) reaction based on mitochondrial cytochrome oxidase I (mtCOI) sequences. A total of 22 SNP haplotypes (i.e., strings of alleles at the four SNP loci) was observed; haplotype diversity (Hd) = 0.811 (s.d. = 0.008; Var = 0.00006). Analysis of molecular variation within and among samples, areas (i.e., Marguerite Bay, Crystal Sound, shelf, offshore), and collection years revealed no difference between 2001 and 2002 collections as a whole, although 2001 and 2002 collections from Marguerite Bay explained 7.4% of the variance (Fsc = 0.072; P = 0.002 $\pm$ 0.001). Most of the variation (96.3%) occurred within samples each year, with no differentiation among areas. There was small, but significant differentiation among samples within areas in 2001 (4.6%; Fsc = 0.045; P = 0.015 ± 0.003) and 2002 (6.3%; Fsc = 0.062; P = 0.000 $\pm$ 0.000). Separate analysis of furcilia larval life stages revealed large and significant variation among samples each year: in 2001, among-sample variation was 25.6% (Fsc = 0.225; 0.095 ± 0.010) for F3 furcilia, 73.0% for F4, 15.8% for F5, 12.2% for F6; in 2002, among-sample variation was 191.% for F1 and 9.2% for F2 stages. For 2001 collections from Marguerite Bay, 20.6% of the variation (Fsc = 0.199, P = 0.000 $\pm$ 0.000) was found among life stages within samples. The significant differentiation

among life stages within and among samples and between years is interpreted to suggest multiple sources of recruitment of *E. superba* in the WAP region, consistent with advective transport and production in both offshore and shelf habitats. In light of climate change and global warming, as well as an expanding commercial krill fishery, further population genetic analysis at sub-regional scales is needed to understand and eventually predict population dynamic processes (e.g., recruitment, migration, retention, and over-wintering) of the Southern Ocean krill.

1.INTRODUCTION

1.1 The Western Antarctic Peninsula region

One of the most biologically-rich and climatologically-vulnerable areas of the Southern Ocean, the Western Antarctic Peninsula (WAP) region is known to support large standing stocks of the Southern Ocean krill, *Euphausia superba* (Marr, 1962; Siegel, 2005; Atkinson et al., 2008). The continental shelf in the WAP region is 200 to 500m deep and is intersected by several depressions (500 to 700m), which provide connections between the shelf and the offshore oceanic waters (Hofmann et al., 1996). This topography drives the circulation patterns, including a large cyclonic gyre with sub-gyres at the northeastern and southwestern ends of the shelf (Stein, 1992, Smith et al., 1999; see Figure 1). Of particular note is Marguerite Bay, where coastal currents are thought to turn and flow counter-clockwise, creating a potential retention mechanism for krill and the associated pelagic community.

The Antarctic Peninsula region is one of the most rapidly warming regions on Earth (Clarke et al., 2007; Ducklow et al., 2007; Loeb et al., 2009). Winter air temperatures are increasing, duration of winter sea ice cover is lessening, and ice concentration is decreasing (Ducklow et al., 2007). These changes may result in changes in the biological community,

including shifts in both phytoplankton (Moline et al., 2004) and zooplankton (Atkinson et al., 2004, 2008) composition and abundance. Furthermore, these changes may have significant biogeochemical implications in terms of shifting the amount of carbon channeled to higher trophic levels versus exported to the benthos (Moline et al., 2004) and affecting dimethylsulfide (DMS) production (Kasamatsu et al., 2004).

1.2 Southern Ocean krill

Euphausia superba is a swarming crustacean, whose discontinuous distribution and highly variable time/space patterns of concentration have long been a subject of research (Marr, 1962; Spiridonov et al., 1996; Lawson et al., 2008). Areas of higher concentration of adult krill are associated mainly with major oceanic gyres in the Weddell and Ross Seas, which are formed by the Antarctic Circumpolar Current, the main surface current of the Antarctic Ocean. This association has suggested that gyres may generate a retention mechanism that promotes the formation of separate, self-supporting stocks of krill (Mackintosh, 1973).

Both reproducing and non-reproducing krill are consistently found in the WAP region and krill may be retained in this region by persistent gyres situated over the northern shelf (Marr, 1962, Lascara et al., 1999). Marguerite Bay and other locations along the Antarctic Peninsula are hypothesized to be important over-wintering grounds for krill, and may also be important source regions for krill populations in the Bransfield Strait around South Georgia (Brinton, 1991; Huntley and Brinton, 1991; Atkinson et al., 2001; Fach et al., 2002).

In the WAP region, Lawson et al. (2008) showed a clear relationship between krill, bathymetry, and the presence of Circumpolar Deep Water (CDW), which may explain the large spatial patchiness of krill as a result of water mass structure along the Antarctic continental shelf. Interactions with sea ice can further increase population retention or dispersal of krill at regional

scales (Thorpe et al., 2007), making spawning and nursery areas of krill in the WAP region, where winter sea ice duration is shortening, particularly sensitive to environmental change (Atkinson et al., 2004).

1.3 Population genetics of Southern Ocean krill

The population genetic diversity and structure of Southern Ocean krill has been extensively studied. For *Euphausia superba* in particular, allozymes were used extensively during the 1980s for population genetic analysis. Overall, population genetic investigations of Antarctic krill populations based on allozyme data have suggested genetic homogeneity for the species. These studies revealed considerable genetic variation within the species' population, with significant sample-to-sample heterogeneity in allelic frequencies (Fevolden, 1986), but – despite continued examination – no evidence of genetic distinctiveness of stocks or geographic populations (Grant, 1983; MacDonald and Schneppenheim, 1983; Schneppenheim and MacDonald, 1983; Kuehl and Schneppenheim, 1986; MacDonald et al., 1986; Marquez, 1987; Fevolden, 1988; Fevolden and Schneppenheim, 1989). Despite an early report of genetic differentiation between two samples collected from each side of the Antarctic Peninsula (Fevolden and Ayala, 1981), further examination of a larger data set, which also included two populations, showed that krill apparently behave as a single breeding unit (Fevolden and Schneppenheim, 1989).

For *E. superba*, Zane et al. (1998) demonstrated significant temporal variation of populations based on DNA sequence analysis of the mitochondrial NADH dehydrogenase (ND-1) gene. The presence of distinct gene pools in the South Georgia area in different years, as well as the genetic differentiation of these samples from those collected in the Weddell Sea, suggested complex population dynamics in the area of the Antarctic Peninsula and surrounding waters. This appears particularly interesting when compared to the observed genetic homogeneity of

geographically distant populations from the Bellingshausen Sea and the Ross Sea (Zane et al., 1998).

Additional molecular genetic analyses of *E. superba* to date have included sequencing of the mitochondrial genome (Machida et al., 2004); phylogenetic analysis using mt16S rRNA (Patarnello et al., 1996); purification of a tRNA (Oshima et al., 1981) and two ribonucleases (Van et al., 1982); chromosomal analysis (Yabu and Kawamura, 1984; Ngan et al., 1990); a study of UVB-induced DNA damage (Malloy et al., 1997); and determination of the DNA barcode region of mitochondrial cytochrome oxidase I (Jarman et al., 2002; Bucklin et al., 2007).

Choice of Population Genetic Markers

Selection of molecular markers for studying the genetic diversity and structure of natural populations is largely dependent on the scale of investigation that one intends to perform: the general rule is that the finer the temporal or spatial scale of investigation, the more polymorphic the molecular marker must be. In this context, there are several genetic markers that can be used for population analysis (see review by Avise, 1994, 2000).

Mitochondrial DNA has become an important resource for studying the history of populations due to the lack of recombination that occurs in the mitochondria. The mitochondrial genome is passed down from mother to offspring completely unchanged except for occasional spontaneous mutations. This unique inheritance pattern makes mitochondrial DNA more susceptible to genetic divergence of geographic populations. Indeed, mitochondrial DNA has proven very useful for studies of population genetic variation and phylogeographic patterns for diverse organisms and various environments, including marine species (e.g., Galindo et al., 2006; Hedgecock et al., 2007). For krill in particular, mitochondrial DNA sequence variation has proven useful to resolve large-scale structure associated with physical barriers to gene flow and –

to a lesser extent – with large scale and persistent ocean circulation patterns (Bucklin et al., 1997, 2002; Zane et al., 1998, 2000; Zane and Patarnello, 2000; Papetti et al., 2005).

In the search for more variable molecular markers to allow resolution of population genetic structure of high gene flow species, recent interest has focused on allelic variation at single nucleotide polymorphisms (SNPs; e.g., Nicholson et al., 2002; Manel et al., 2003). SNPs occur when a single nucleotide (A, T, C, or G) in the gene sequence is altered. SNPs occur throughout the genome and are the most frequent and widespread type of variation in DNA (Wang et al., 1998). SNPs at four-fold degenerate sites in coding regions provide highly variable markers for detection of population genetic diversity and structure. SNP identification, screening, and analysis are increasingly rapid and inexpensive, through creation of relatively inexpensive SNP detection kits and standard protocols (Qintáns et al., 2004). The availability of commercial kits opens up new opportunities for analysis of large samples of zooplankton species with enormous population sizes. SNPs in mitochondrial genes may be inherited as haplotype blocks (and analyzed accordingly), while SNPs in nuclear protein-coding genes offer the advantage of diploid (but unphased) genotypes for population genetic characterization.

The goals of this study are to examine population genetic diversity and structure of the Southern Ocean krill *Euphausia superba* in the WAP region, based upon allele frequencies at mitochondrial SNP loci. The highly variable molecular markers have been selected in order to allow examination of genetic differentiation among samples and areas within the WAP, at spatial scales that are smaller than those of previous studies of this species. In particular, we seek to examine population genetic data in light of hypotheses regarding patterns of advective transport and recruitment of krill in WAP coastal areas and adjacent shelf regions.

2.METHODS

2.1 Sample collection

Zooplankton samples for this project were collected during two cruises of the R/V/IB *Nathaniel B. Palmer* associated with the US Southern Ocean GLOBEC program: NBP-0103 (April 23 - June 6, 2001) and NBP-0202 (April 9 – May 19, 2002). Comprehensive sampling during the cruises was done with the MOCNESS (Multiple Opening Closing Net and Environmental Sensing System; Wiebe et al. 1985), with additional collection and analysis of bioacoustical, hydrographic, and other oceanographic data and observations.

Samples for genetic analysis were collected during the integrated down-haul of 16 MOCNESS tows (Figure 2) and preserved immediately after collection in 95% ethyl alcohol or flash-frozen in liquid nitrogen. Samples were selected for analysis based on comparison of the MOCNESS sampling locations to the bioacoustical characterization of krill patch structure for these same cruises by Lawson et al. (2008), with the goal of selecting samples that fell within one of the two or three "hotspots" (i.e., areas with many aggregations) observed during each cruise.

Samples were also selected to represent one of four defined areas: shelf and Marguerite Bay (sampled in 2001 and 2002), offshore (2001), and Crystal Sound (2002; see Table 1). The areas were defined based on the circulation patterns observed in the WAP region and from published descriptions of current flow (Hofmann et al., 1996; Klinck at al., 2004).

A total of 565 specimens of *E. superba* were sorted from the 16 selected samples. Individuals were sized and identified to life stage at the time of identification. For all samples except four with very large numbers of krill (MOC 6, 21, 22 ,23), all identifiable *E. superba* were removed for analysis.

2.2 Molecular analysis

2.2.1. Design of species-specific primers and SS-PCR

A multiplexed, species-specific PCR (SS-PCR) reaction was designed to discriminate three euphausiid species occurring in the WAP, *E. superba, E. crystallorophias*, and *Thysanoessa macrura*, based on DNA sequence variation of mitochondrial cytochrome oxidase I (COI). A ~700bp region of mtCOI was amplified using consensus primers from Folmer et al. (1994) and Bucklin et al. (2007). Species-specific primers were designed according to Bucklin (2000) and Hill et al. (2001). Conserved regions of 20-30 base-pairs that were diagnostic of each species were identified for each species (Table 2). Priming sites were separated by ~100 base-pairs to allow multiplexing of the reaction. The reaction was carried out by adding all three species-specific primers and an anchor or common primer (Gibbs et al., 1981). PCR reaction parameters were: 94°C for 0.5 min; 63°C for 0.5 min; 72°C for 3 min for 35 cycles. Identification of the species was based upon the size of the PCR product, as detected by agarose gel electrophoresis.

2.2.2. SNP identification and primer design

DNA sequences for a 1,200 base-pair region of the mitochondrial gene, cytochrome b (CytB), were determined using species-specific PCR primers with the following sequence:

Cytb-53F CAC TGG TTG ACC TTC CTG CT

Cytb-1025R CGA GCT CCG ATC CAT GTT AG.

Sequences were determined for 50 individuals selected to maximize the temporal and spatial coverage across the sample set. Single nucleotide polymorphisms (SNPs) were identified based on the pattern of variation in the coding region of the gene; sites of silent substitutions (i.e., nucleotide substitution changes the DNA sequence, but not the amino acid sequence) and especially 3- and 4-fold degenerate sites were chosen. In order to allow design of primers for the

single-base extension detection system, SNP sites selected were adjacent to conserved regions of 20 – 30 base-pairs.

Four SNPs in the mtCYB gene were selected that showed high rates of substitution, multiple alleles, and conserved adjacent sequences (Table 2). All SNPs were silent substitutions at three- or four-fold degenerate sites in the third codon position of the nucleotide sequence. The SNP sites and primers were named according to their position in the *E. superba* mitochondrial genome sequence (GenBank Acc. No. AB084378; Machida et al., 2004).

SNP alleles were detected using the Applied Biosystems, Inc. (Foster City, CA) SNaPshot kit. The commercial kit is based on a multiplexed single-base extension reaction (Linblad-Toh et al., 2000). Primers were designed using Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3) and tested for hairpin structures and self-compatibility using ABI Primer Express software. The NCBI BLAST tool (http://www.ncbi.nlm.nih.gov) was used to check for other sequence homologies. Primers were synthesized with added runs of "A" or "T" to create different length products (ranging from 25 to 53 base-pairs) to allow accurate detection in multiplexed reactions (Table 2).

2.2.3. Detection of SNP allele frequencies

Genomic DNA was extracted from excised muscle tissue of individual krill using a commercial kit (DNAeasy, Qiagen, Valencia, CA). The whole specimen was used for analysis of smaller individuals. Before analysis, the DNA was tested to confirm species identification using a multiplexed, species-specific PCR (SS-PCR) reaction.

DNA from each individual was used to amplify the 1,200 base-pair region of CytB. PCR products were purified with the UltraClean PCR Clean-Up Kit (Mo Bio, Carlsbad, CA) and used for the SNaPshot minisequencing reaction. The single-base extension reaction was carried out in

a total volume of 10 μl comprising 5μl of the SNaPshot buffer, 4μl PCR product, 1μl of extension primers mix (final concentration 2μM). The reaction was performed in an ABI GeneAmp 9700 machine, with 54^{o}C annealing temperature, and reaction protocols as specified in the kit documentation. The SNaPshot assay was detected using an ABI 3130 Genetic Analyzer and analyzed with ABI Gene Mapper software.

2.3 Data Analysis

2.3.1. Diversity and neutrality

Alleles were detected at the four CytB SNP sites for all individuals; allele frequencies were determined for samples and areas in 2001 and 2002. Analysis was performed for data from 282 individuals collected in 2001 and 283 in 2002. Metrics of genetic diversity, including haplotype and nucleotide diversity, were calculated using DNAsp software, Ver. 4.10 (Rozas et al., 2003).

2.3.2 Analysis of population structure

Analysis of SNP allele frequencies was carried out using AMOVA, Version 3.11 (Analysis of MOlecular Variation; Excoffier et al., 2005) to analyze SNP haplotype (i.e., the string of alleles at the 4 SNP sites) frequency and distance. Analysis of F-statistics (Weir and Hill, 2002) and variance partitioning was done within and among samples, areas (as defined in Table 1), and years. In each case 1000 pseudoreplicate datasets were created from the original dataset to use as a null distribution for significance testing.

Tests for significant departures from selective neutrality of the haplotypes in each population were executed with Arlequin version 3.11 (Excoffier, Smouse et al. 1992). Analysis of population structure was done using DNAsp to visualize the geographic distribution and genetic variation of the SNP haplotypes. Network analysis was carried out using the median joining algorithm (Bandelt et al., 1999).

3.RESULTS

3.1 Specimens analyzed

The 565 individuals of *Euphausia superba* were identified and confirmed by SS-PCR, including 275 individuals from nine samples collected in 2001, and 281 individuals from seven samples collected in 2002 (Table 1). All 6 furcilia larval stages, juveniles, and adults were among the specimens selected for analysis (Table 3; Figure 3). Considering all samples and both years, all life stages were fairly evenly represented.

3.2 Population Genetic Diversity

All four CytB SNP sites showed high levels of variation, with four (CYB-519), three (CYB-681 and -789), or two (CYB-915) alleles (Figure 4). Metrics of genetic diversity were calculated: haplotype diversity, Hd = 0.811 (standard deviation = 0.008; variance = 0.00006); nucleotide diversity, $\pi = 0.438$; and coancestry coefficient, $\theta = 0.289$. The effective number of alleles, k = 1.751.

A total of 22 "haplotypes" (i.e., strings of alleles at the four SNP loci) were observed among the 565 individuals of *Euphausia superba* collected in the WAP region during 2001 and 2002. Of these, six haplotypes occurred in both 2001 and 2002; two haplotypes were found only in 2001; and five haplotypes were found only in 2002 (Figure 5). In general, haplotypes observed only in one of the two years were less frequent than those observed in both years.

3.3 Population Genetic Structure

Comparison of SNP allele frequencies for all samples each year revealed no evidence of differences between the 2001 and 2002 collections (Fct = 0.00; percent variation = -0.21; P = 0.353+0.017; Table 4), except the Marguerite Bay collections in 2001 and 2002 showed

significant genetic differentiation, explaining 7.4% of the variance (Fsc = 0.072,P = 0.002 ± 0.001).

Evaluation of SNP allele frequencies among *E. superba* collected in different WAP areas – defined based on bathymetry and circulation patterns (see Figure 1) and analyzed by grouping samples accordingly (see Table 1) – revealed no genetic differentiation among areas in either 2001 (Shelf, Marguerite Bay, and Offshore) or 2002 (Shelf, Marguerite Bay, and Crystal Sound). There was a small-but-significant level of genetic differentiation among samples within the areas in both 2001 (4.6%; Fsc = 0.045; P = 0.015 ± 0.003) and 2002 (6.3%; Fsc = 0.062; P = 0.000 ± 0.000), while most of the variation was found with samples (97.3% in 2001; 95.8% in 2002; Table 5).

When some furcilia larval life stages were considered separately, a larger fraction of the variation (9.23% – 25.6%) was explained by comparisons among samples within the defined areas. In 2001, separate analysis of F3 – F6 furcilia larval life stages revealed large and significant variation among samples within areas, ranging from 72.96% (Fsc = 0.426; 0.066 ± 0.006) for F4 to 15.8% (Fsc = 0.137; P = 0.067 ± 0.006) for F6 (Table 6). In 2002, large variation among samples within areas was observed for furcilia larva stages F1 (18.07%; Fsc = 0.160; P = 0.000 ± 0.000) and F2 (9.23%; Fsc = 0.080; P = 0.001 ± 0.001). In all cases, within-sample variation predominated (Table 6). Comparison of adult *E. superba* collected from Crystal Sound and Marguerite Bay in 2002 explained 12.4% of variation (Fsc = 0.124; P = 0.318 ± 0.013). Among samples collected in 2001 from Marguerite Bay, a large fraction of the variation (20.6%) was explained by variation among life stages within a sample (Fsc = 0.199, P = 0.000 ± 0.000).

4.DISCUSSION

The decades-long quest to fully and accurately characterize patterns of population genetic diversity and structure of the Southern Ocean krill *Euphausia superba* has left many issues unresolved. From early studies with allozymes (e.g., Ayala and Valentine, 1979; Fevolden, 1984) to more recent studies with molecular markers (e.g., Zane et al., 1998), these studies have revealed high levels of genetic diversity, with little evidence of significant temporal or spatial genetic structure, except for some findings of sample-to-sample variation (e.g., Fevolden, 1986).

This study is designed to bring to bear highly-variable molecular markers, allelic variation at sites of single nucleotide polymorphism (SNP) in a mitochondrial gene, to examine genetic diversity and structure of *E. superba* within the Western Antarctic Peninsula (WAP) region. The focus on krill populations at the sub-regional scale is motivated by the goals of the US GLOBEC Southern Ocean program, which seeks to understand krill population dynamics, including time/space patterns of reproduction and recruitment, in the oceanographically and bathymetrically complex WAP region (see Hoffman et al., 2004; 2008).

4.1 Krill recruitment hypotheses in the WAP region

The spawning season of krill is limited to the Antarctic summer, during which *Euphausia superba* is thought to complete development from egg through naupliar, calyptosis, and furcilia larval stages (Marr, 1962). Adult *Euphausia superba* are known to spawn in deep water (>500 m), releasing eggs at depth that develop through larval stages as they rise to the surface (Smith and Schnack-Schiel, 1990). In the WAP region, krill are believed to spawn in deep offshore waters in the Circumpolar Deep Water (CDW; Hofmann et al., 1992; Hofmann and Klinck, 1998; Siegel, 1988, 2005; Ross et al., 2000; Atkinson et al., 2001; Fach et al., 2002). Spawning may also occur in deep basins and troughs on the shelf (Brinton, 1991), with retention of larvae

and juveniles in coastal waters, such as Marguerite Bay and Crystal Sound, which have retentive current pathways (Klinck et al., 2004).

Previous work in the WAP region has suggested that adult *E. superba* populations may be concentrated in off-shelf habitats during summer, migrating to more coastal waters during winter periods of increased ice cover (Siegel, 1988; Lascara et al., 1999). Siegel (1988) also observed that krill abundance increases after the Spring retreat of the pack ice, when the krill may move into open waters. Larger krill may form dense aggregations at any time of year (Kawaguchi et al., 1986; Gutt and Siegel, 1994; Lascara et al., 1999) and may overwinter in the water column (Ross et al., 1996).

During fall 2001, larval krill were dispersed over the entire width of the shelf, with greatest abundances seen along the shelf-break (Ashjian et al., 2008). Wiebe et al. (this volume) reported the broad distribution of *E. superba* larvae in the WAP region during 2001 and 2002; nine of the 16 samples analyzed for this study were also used by that study for determination of life stage frequency distributions. Our life stage counts (Table 3; Figure 3) do not differ from those of Wiebe et al., allowing us to conclude that our analysis is representative of the populations sampled.

Questions of the sources of recruitment and pathways of dispersal for *E. superba* in the WAP region are not amenable to population genetic analysis. Even for highly variable molecular markers, the level of advective mixing and dispersal (gene flow) will be expected to prevent the genetic divergence between geographic populations necessary to infer recruitment sources. Thus, this study cannot suggest a possible source, either offshore or coastal, of recruitment of larvae to the WAP region.

The finding of small, but significant differentiation among samples within an area or region (i.e., meso- to large spatial scales; see Table 5) is consistent with previous studies of *E. superba* (Fevolden, 1986; Zane et al., 1998) and other euphausiids (Bucklin and Wiebe, 1986; Bucklin et al., 1997, 2002; Zane et al., 2000; Papetti et al., 2005). Also typical of euphausiids – and other zooplankton – is the lack of significant variation explained by comparisons among groups of samples in the same oceanographic or bathymetric domain (i.e., referred to as areas in this study; see Table 5). The consistent finding of small but significant differentiation in time and/or space of krill populations may result from the difficulty of detecting population differentiation for species – such as *E. superba* – with enormous estimated effective population size (i.e., the number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift). It is likely that this study was able to detect only a fraction of any genetic structure due to the sample sizes analyzed.

Of particular relevance for examination of recruitment process in *E. superba* are two population genetic results from this study. First is the large and significant differentiation among samples for furcilia larval stages F3 – F6 in 2001 and F1 – F2 in 2002 (Table 6). Since the molecular markers are silent substitutions in the coding region of CytB, such variation would not seem to result from differential mortality (natural selection) during cohort maturation. We speculate that this pattern of variation may result from cohorts of larvae that are advected from diverse spawning areas throughout the WAP, and adjacent areas and beyond. A second finding with relevance to patterns of recruitment is significant differentiation among life stages in samples collected from Marguerite Bay in 2001 (20.6%; Fsc = 0.199; P = 0.000 $\pm$ 0.000). This collection of 102 individuals from 5 samples (including all 6 furcilia stages, juveniles, and adults; see Table 3) was made in a krill patch with very high *E. superba* concentration (see

Wiebe et al., 2004). The genetic heterogeneity of the different life stages again suggests multiple recruitment events and sources, with a more complex pattern of recruitment than that of a single, panmictic population. However, caution is needed in the interpretation of these results, given the small numbers of individuals of each life stage analyzed per samples.

The sub-regional scale patterns of genetic diversity and structure of *E. superba* suggest geographically distributed reproduction and recruitment processes. Thus, *E. superba* may spawn in deep waters of the WAP region (e.g., Hofmann et al., 1992; Ashjian, 2008), with transport of larval krill onto the shelf (Capella et al., 1992). Additionally, the krill may spawn in coastal regions, with retention of eggs through larval and juvenile development (Brinton, 199?, Quetin and Ross, 2001; Wiebe et al., this volume). The genetic differentiation of the same furcilia larval stage among different samples and among different furcilia stages in the same sample suggest that recruitment and dispersal patterns in the WAP may vary over time. Loeb et al. (2009) suggest that sustained positive values of the Southern Oscillation Index (SOI) may promote exchange between the coastal and offshore krill populations in the WAP region. Interestingly, values of the SOI index were positive for both 2001 and 2002, suggesting enhanced cross-shelf mixing of the *E. superba* population. Such temporal variation is consistent with the observed patterns of population genetic diversity and structure within and among life stages and samples of *E. superba* in the WAP region.

4.2 Species identification using SS-PCR

The discrimination of furcilia larva life stages of *E. superba* for this study relied upon the design of a PCR-based approach to species identification using species-specific oligonucleotide primers designed from the mitochondrial cytochrome oxidase I (COI) gene region (see Bucklin et al., 2007). The multiplexed, species-specific PCR (SS-PCR; see Bucklin, 2000) reaction

discriminated three euphausiids that co-occur in the WAP region: *Euphausia superba, E. crystallorophias*, and *Thysanoessa macrura*. The SS-PCR screening, which was done for all specimens used in this study, is particularly important and useful for ensuring accurate identification of early life stages and of specimens preserved in alcohol, which preserves the DNA, but may allow damage and loss of diagnostic characters.

5. CONCLUSIONS

Patterns of genetic diversity and structure of the Southern Ocean krill, *Euphausia superba,* were described based on analysis of allele frequencies at four sites of single nucleotide polymorphisms (SNPs) in the mitochondrial cytochrome b (CytB) gene. Collections were made during US GLOBEC Southern Ocean cruises to the Western Antarctic Peninsula (WAP) region in 2001 and 2002. A total of 565 *E. superba* (including all 6 furcilia larval stages, juveniles, and adults) were identified to stage, with species identifications confirmed using a multiplexed, species-specific PCR (SS-PCR) reaction. Analysis of molecular variation revealed no significant differentiation between 2001 and 2002 samples considered together, although 2001 and 2002 collections from Marguerite Bay showed small but significant differentiation. While most of the variation was found with samples (97.3% in 2001; 95.8% in 2002), there was significant variation among samples within WAP areas in both years. Larger levels of differentiation were observed among samples when furcilia larval stages were analyzed separately, and among life stages in the 2001 collection from Marguerite Bay. We conclude that a possible explanation for the observed patterns of population genetic diversity and structure of *E. superba* populations in the WAP region is multiple centers of reproduction and recruitment, utilizing coastal, shelf, and offshore areas. Further use of allele frequencies at SNP sites in mitochondrial genes is warranted for

studies of population genetic diversity and structure of high gene flow species at regional to sub-regional spatial scales.

6. **ACKNOWLEDGEMENTS**

We thank the Captains and crew of the RVIB *Nathaniel B. Palmer* for their contributions to this effort. We acknowledge the assistance of Gareth Lawson (Woods Hole Oceanographic Institution) in sharing the early results of his krill patch studies in the Western Antarctic Peninsula region. Technical support and assistance was provided by Lisa M. Nigro (University of Connecticut) and Erica Bortolotto (University of Padova). Funding was provided by NSF Office of Polar Programs (Award Nos. OPP-0338195 and OPP-0610789 to A.B. and T.P. and OPP-0431357 to P.H.W.). Additional funding was provided to T.P. by the Italian National Antarctic Research Programme.

Figure Legends:

Figure 1. Overview of the Western Antarctic Peninsula general circulation pattern. Major features include the following: the Antarctic Circumpolar Current (ACC) running to the north and east beyond the shelf break; the Antarctic Peninsula Coastal Current (APCC) running along the coastline to the southwest (during ice- free conditions); clockwise gyres on the mid-to outer shelf south of Bransfield Strait; and Circumpolar Deep Water (CDW) coming on to the shelf at specific locations along the peninsula. Figure re-drawn from one provided by R.Beardsley, C. Moffat, M. Zhou et al.; sources include and Moffata et al. (2008).

Figure 2. Collection locations for *E. superba* used in this study : (A) Samples collected during 2001 (NBP-0103). (B) Samples collected during 2002 (NBP-0202). Samples are labeled using the number of the MOCNESS tow. Shading indicates areas as defined for population genetic analysis. See Table 1 for collection information and text for further explanation.

Figure 3. Distribution of life stages for samples of *E. superba* collected during 2001 (above) and 2002 (below). MOCNESS tow number in the legend indicates samples; shading of the histogram bars indicates areas: Shelf (gray shaded bars) Marguerite Bay (cross-hatched) for both 2001 and 2002; Offshore (black in 2001); and Crystal Sound (black in 2002).

Figure 4. Allele frequencies at four SNP sites in the mitochondrial cytochrome b gene of *E. superba* for samples collected in 2001 (above) and 2002 (below). Allele frequencies were calculated for samples grouped by area, as described in the text. Area abbreviations are: Marguerite Bay (MB), Offshore (OS), Shelf (SH) and Crystal Sound (CS). Allele names are nucleotide bases: Adenine (A), Cytocine (C), Guanine (G), and Thymine (T). Missing alleles were not found in any analyzed sample.

Figure 5. Network diagram showing frequencies, relationships, and source locations for the 22 SNP haplotypes observed for *E. superba* collected during 2001 and 2002. The four most frequent haplotypes (H06, H08, H18, H19, and H21) were found in both 2001 and 2002. Relative frequencies are indicated by the size of the pie; colors of the slices indicate the collection area and year for that individual (see legend).

Table 1
Collection information for *E. superba* used for population genetic analysis from Southern Ocean GLOBEC cruises in 2001 and 2002. Samples are indicated by MOCNESS tow number (Tow), as shown on the map of the collection area (see Fig. 2). Numbers of individuals analyzed for SNP variation (*N*).

Area	Tow	Date	Local time	Depth (m)	Latitude	Longitude	*N*
2001							
Shelf	3	5/2/2001	0320	500	66°48.56	70°22.76′	99
Shelf	12	5/13/2001	1730	350	68°23.7	72° 18.44′	27
Shelf	18	5/20/2001	1155	460	69°31.27	76°17.54′	2
Marguerite Bay	9	5/7/2001	1654	400	68°28.70	68°46.23′	5
Marguerite Bay	21	5/28/2001	0750	100	67°52.90	68°06.15′	6
Marguerite Bay	22	5/28/2001	2202	90	67°52.62	68°26.00′	75
Marguerite Bay	23	5/29/2001	0107	90	67°55.48	68°21.55′	4
Marguerite Bay	24	5/29/2001	0300	600	67°55.30	68°30.69′	7
Offshelf	6	5/4/2001	1332	1000	66°40.16	73°22.08′	50
						Subtotal	**275**
2002							
Shelf	2	4/15/2002	0220	325	66°11.10	69°08.10′	58
Shelf	4	4/18/2002	0456	500	66°46.06	70°10.51′	11
Shelf	16	5/1/2002	2200	400	68°24.24	72°18.24′	86
Shelf	22	5/10/2002	1254	282	69°31.99	76°17.30′	5
Marguerite Bay	9	4/23/2002	0035	800	67°55.3	68°30.13′	30
Marguerite Bay	10	4/24/2002	1340	580	68°28.8	68°47.20′	18
Crystal Sound	24	5/15/2002	0425	195	66°31.29	67°38.50′	73
						Subtotal	**281**

Table 2
DNA sequences for the oligonucleotide primers used in this study. **Above**: Species-specific primers used in the SS-PCR reaction to discriminate three co-occurring krill species and confirm identification of *E. superba* at all life stages. Primer names indicate the species and the size of the amplified product (in numbers of nucleotide base-pairs): *E. superba* (Es-594), *E. crystallorophias* (Ec-460), and *Thysanoessa macrura* (Tm-330). The SS-PCR reaction also included a conserved or "anchor" primer (COI-164F). **Below**: Primers for the multiplexed single-base extension reaction for detection of alleles at the four single nucleotide polymorphism (SNP) sites analyzed in this study. Added strings of "T" or "G" preceding the species-specific detection region allow multiplexing of the SNP reactions (see text for explanation). SNP sites were designated by the base position in the coding region of the gene; primer sequences are shown in the 5′-to-3′ direction.

SS-PCR	**Primer sequence**
Es-594	GAGAGAGAGGAGAAGTAGGATAGCTGTAATAAAC
Ec-460	AAGAGTTAAGGAAGGAGGCAGTAACC
Tm-330	CAGCAATATGGAGAGAAAAGATACCC
COI-164F	GTTGTAGTTACAGCACATGCTTTTGTTATAA
SNPs	
CytB-519	GGAGGGTTTGCTGTAGATAA
CytB-681	GGGGCATTCCACCATTATTTTACTCT
CytB-789	TTTTTTTTTTTTTTTTCCTGCCAACCCTTT
CytB-915	GGGGGGGGGGATCAATCGCTATTTTATTAATTTT

Table 3
Numbers of individuals of *E. superba* analyzed for SNP allele frequencies from collections during 2001 and 2002 cruises to the Western Antarctic Peninsula region. Total numbers are indicated for each life stage and each area (as explained in the text). Life stages include the six furcilia larval stages (F1–F6), juvenile (Juv) and adult (Ad).

Area	F1	F2	F3	F4	F5	F6	Juv	Ad	Total
2001									
Marguerite Bay	1	1	0	8	37	46	1	3	97
Shelf	36	0	27	21	14	30	0	0	128
Offshore	12	3	10	15	9	1	0	0	50
2002									
Marguerite Bay	3	4	1	10	0	0	30	0	48
Shelf	44	57	43	6	0	3	5	2	160
Crystal Sound	0	0	0	0	0	0	5	68	73
Total	96	65	81	60	60	81	41	73	556

Table 4
Evaluation of interannual variation in SNP allele frequencies for samples of *E. superba* collected during 2001 and 2002 in the Western Antarctic Peninsula region. Results are shown for hierarchical variance partitioning by the Analysis of Molecular Variation (AMOVA; Excoffier eta al., 2002) and include variance explained by comparisons within samples, between samples within each year, and between years. Results are given as degrees of freedom (d.f.); *F*-statistics (F_{sa} , F_{ay} , F_{yt}); percentage of variation [Var(%)]; significance value ± variance from 1000 permutations (*P*-value). Significance levels are indicated by asterisks: $\leq$ 0.1 (*), $\leq$ 0.001 (**), $\leq$ 0.0001 (***), not significant (NS).

	d.f	*F*-statistic	Var (%)	P-value
Within samples (in an area)	540	Fsa=0.036	96.40	0.000±0.000***
Among areas within years	14	Fay=0.038	3.77	0.000±0.000***
Among years	1	Fyt=0.000	0.00	0.321±0.016 NS

Table 5
Evaluation of spatial population genetic structure within and among samples of *E. superba* and areas of the WAP region for 2001 and 2002 based on SNP allele frequencies. Area designations are as described in Table 1 and the text; analysis includes Shelf, Marguerite Bay, and Offshore (for 2001); Shelf, Marguerite Bay, and Crystal Sound (for 2002). See legend for Table 4 for explanation of abbreviations and statistical results.

	d.f	*F*-statistic	Var (%)	P-value
2001				
Within sample	266	Fsa=0.023	97.66	0.004±0.002**
Among samples within areas	6	Fay=0.043	4.36	0.023±0.005**
Among areas	2	Fyt=0.000	0.00	0.664±0.015 NS
2002				
Within sample	274	Fsa=0.041	95.87	0.000±0.000***
Among samples within areas	6	Fay=0.064	6.62	0.000±0.000***
Among areas	2	Fyt=0.000	0.00	0.792±0.012 NS

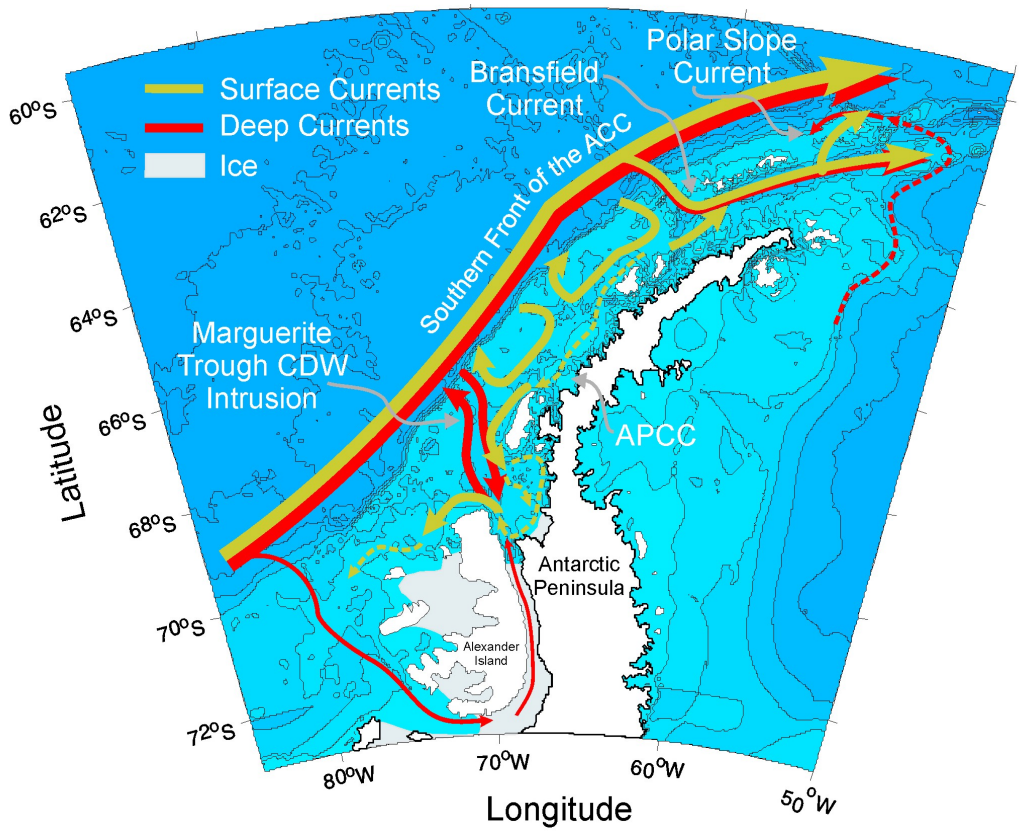


Figure 1.Western Antarctic Peninsula circulation pattern.

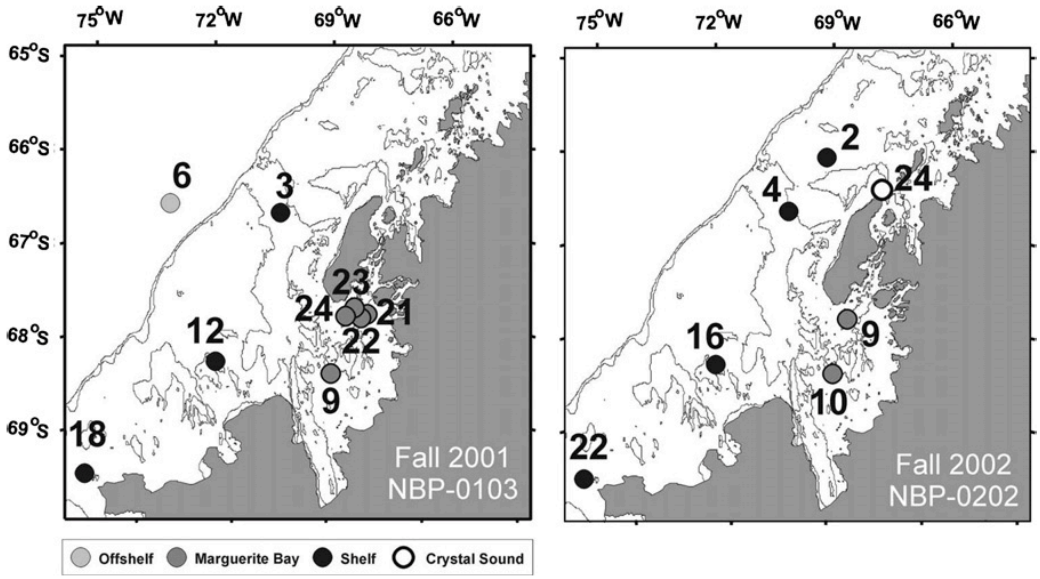


Figure 2. Collection locations for *E.superba*

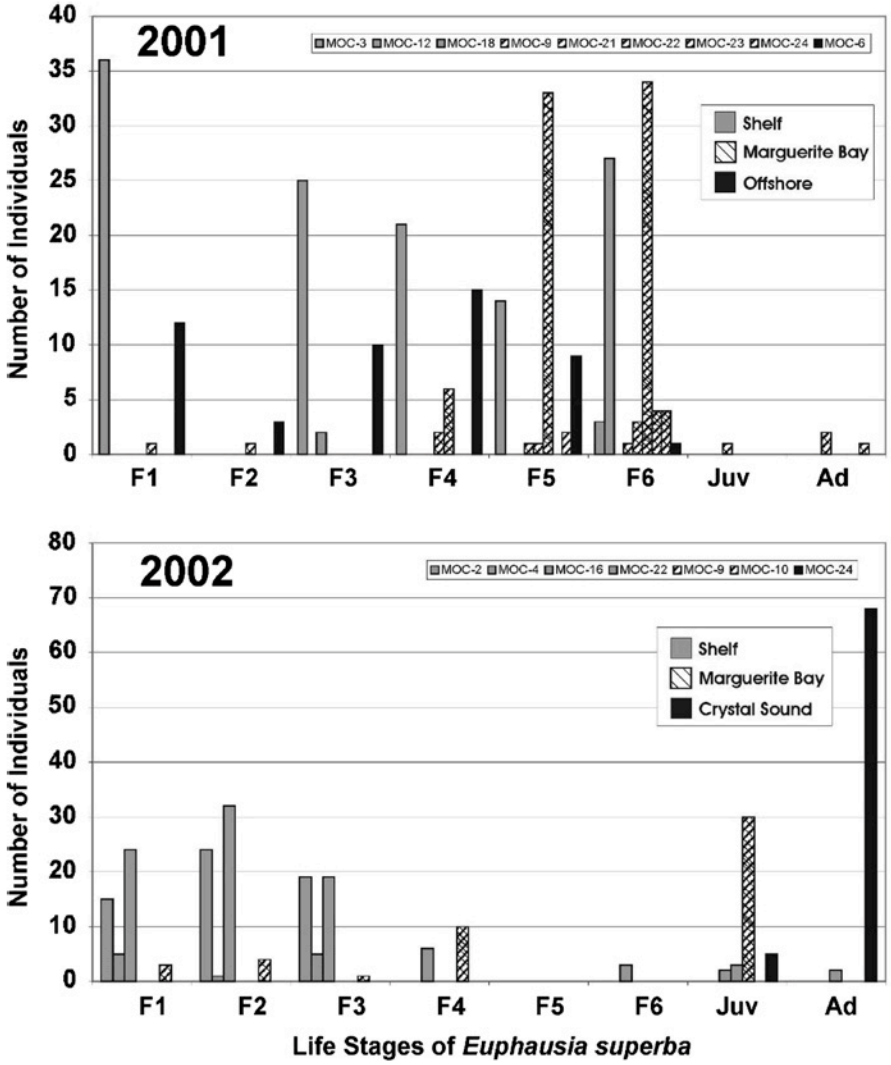


Figure 3. Life stage distribution of *E. superba*

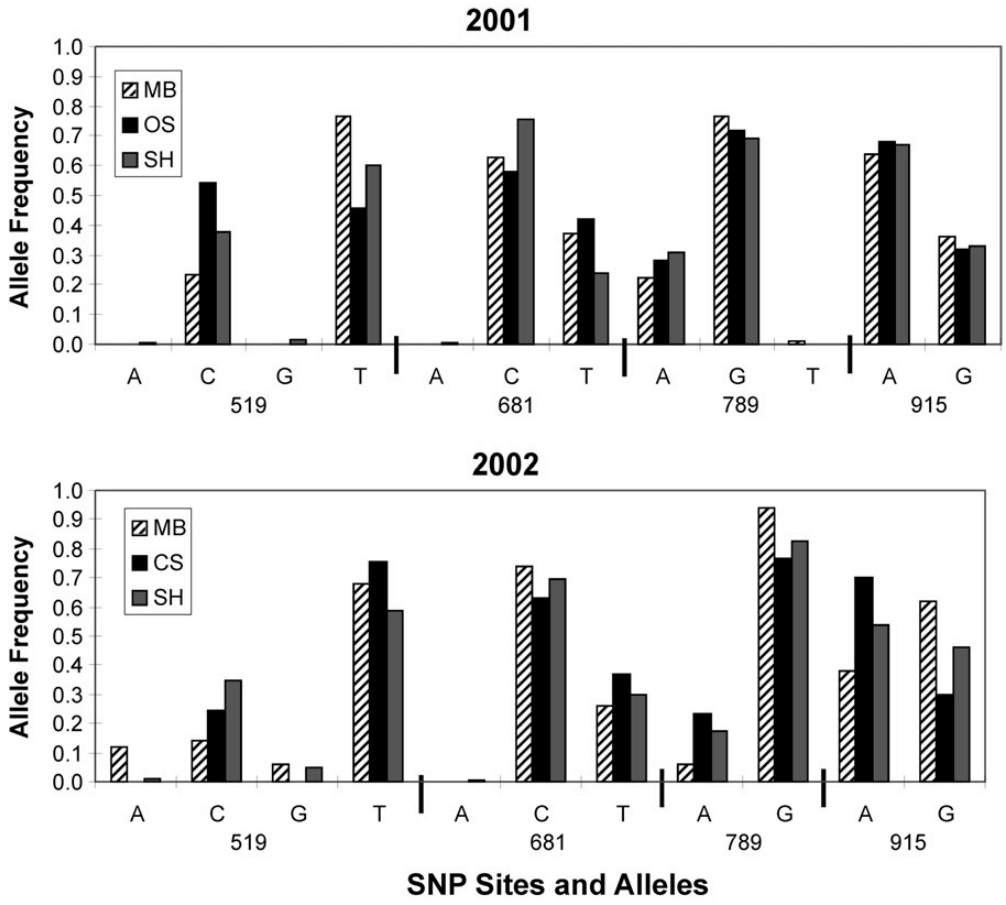


Figure 4. Allele frequencies at four SNP sites

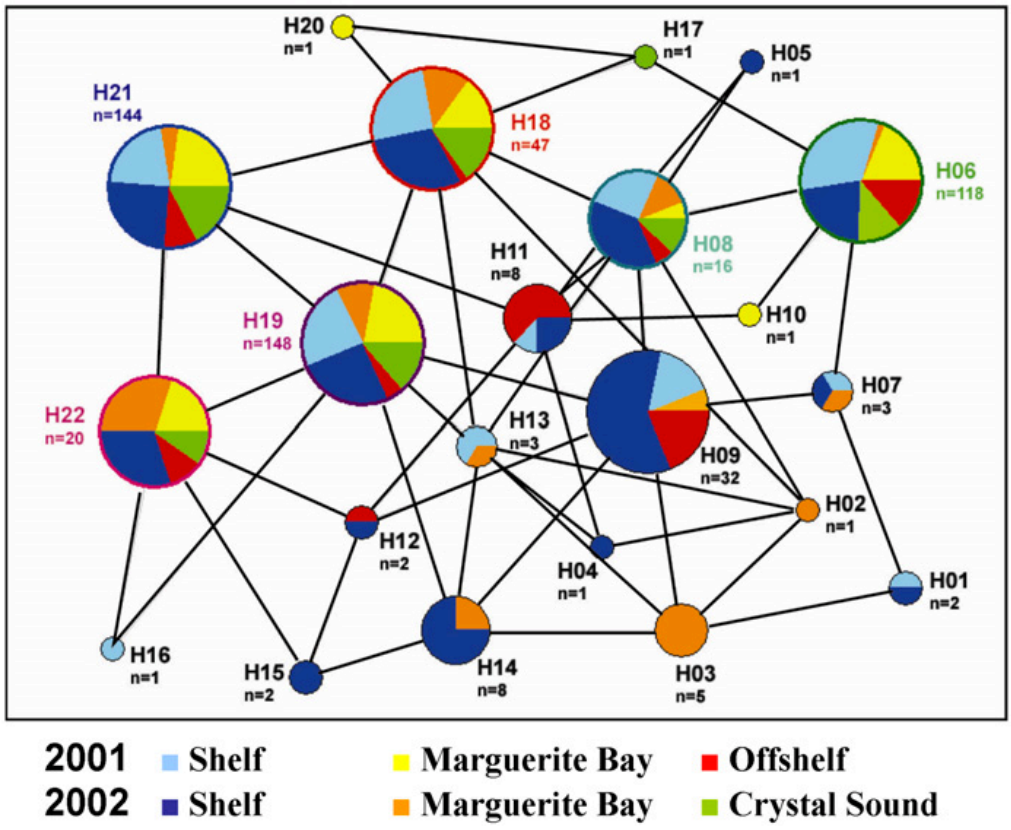


Figure 5.Network diagram of 22 SNP haplotypes for *E.superba*

Summary:

This study took place within the well-studied waters of the Western Antarctic Peninsula (WAP) region (Hofmann et al. 1996; Thompson et al., 2009) and the Indian Sector of the Southern Ocean (Sparrow et al., 1996; Ohshima et al., 2000). The environmental data presented here showed the seasonal differences between spring and summer in the WAP region, including the characteristic seasonal modification of these water masses. Water mass features and distribution in the sampled areas were consistent with previous studies describing the presence and seasonality of this water masses in these regions.

Salps collected during the austral spring of 2011 were associated with Winter Water (WW) and Circumpolar Deep Water (CDW) in the offshore station and with WW and Bransfield Strait Basin Bottom Water (BS) in the Bransfield Strait stations. During austral summer 2011 in the WAP region, salp presence was associated with WW and CDW on the offshore and shelf stations. Similarly, during the austral summer of 2009 in the Indian Sector, salps were present at stations where WW and CDW water masses occurred in the offshore station, while the Coastal Current (CC) influenced more coastal stations.

Based on long-term studies, it has been suggested that nutrients might also have a role affecting zooplankton distribution with the ongoing climate change (Lee 2010). Nutrient distribution was similar in all stations, regions and seasons. Our PCA analysis for austral spring and summer of 2011 in the WAP and austral summer of 2009 in the Indian Sector showed that the main factors influencing *S. thompsoni* distribution were silicate (PC1) and temperature (PC2). However when chlorophyll *a* was included in the PCA analyses for austral spring 2011 in the WAP and austral summer 2009 in the Indian Sector, the main factors influencing distribution of this species were

chlorophyll *a* and temperature (PC1) and silicate (PC2). PCA analysis included in this work contributes to an understanding of *S. thompsoni's* distribution in relation to different environmental parameters (including nutrients).

This work also highlights the differences in life stage distribution in the WAP region during the austral spring and austral summer, as well as the difference in *S. thompsoni* life stage distribution during the austral summer in two different regions of the Southern Ocean: the WAP and Indian Sector. Our field observations and collections from two Southern Ocean regions during two years (2009 and 2011) and both austral spring and summer provide additional understanding and identification of the different environmental conditions in which *S. thompsoni* can be found. Characterization of the environmental conditions associated with salp presence can help understand habitat preferences of the species and thus provide a foundation for predictions of responses to changes in environmental conditions associated with climate change.

Comparatively few studies have used molecular approaches to understand how organisms may respond and adapt to environmental variability at appropriate time and space scales ranging from short-term events to long-term climate change.

Genome annotation for *S. thompsoni* matched sequences described for marine model species such as *Ciona intestinalis* and *Oikopleura dioica* (Tunicata), *Strongylocentrotus purpuratus* (Echinozoa) and *Branchiostoma floridae* (Cephalochordata). Even though both tunicate model species were represented, the *Oikopleura* genome was not among the top 10 species, this can be due to the fact that *Oikopleura* has the smallest genome ever found in a chordate (51 – 65 Mb) (Seo, 2001); on the other hand *Ciona* has a larger genome and has deeper coverage. Therefore, the database will include more genes from *Ciona* to map to, putting this specie on the first place of the top hit species list.

With the use of whole-genome analysis, this study provides a catalog to identify genes that can be used to study how organisms respond and adapt to environmental variability at the molecular level, ranging from short-term events to long-term climate change. Additional genomic analysis should focus on genes with particular biological functions that are relevant to adaptation of *S. thompsoni* to changes in the Antarctic environment, such as temperature and food availability. The number of samples and the sequencing effort produced a complete reference transcriptome, which allowed statistical analysis of transcriptome-wide patterns of gene expression among salps collected in different environmental conditions. This provides a novel analytical approach to understanding how an important Antarctic species may be affected by environmental variation. Future research should involve looking at the community response including krill and other key Antarctic species to fully understand the present and future effects of the changing Antarctic environment on the ecosystem. Additional transcriptomic analysis should focus on genes with particular biological functions that are relevant to adaptation of *S. thompsoni* to changes in the Antarctic environment, such as temperature and food availability. Targeted sequencing could be used to find hundreds of loci for population genetic analysis, with scoring of SNPs from 300-400 genes with functional significance and to provide sensitive markers of population structure and detect selection.

Lastly, patterns of genetic diversity and structure of the Antarctic krill, *Euphausia superba,* were described based on analysis of allele frequencies at four sites of single nucleotide polymorphisms (SNPs) in the mitochondrial cytochrome b (CytB) gene. Collections were made during US GLOBEC Southern Ocean cruises to the Western Antarctic Peninsula (WAP) region in 2001 and 2002. We conclude that a possible explanation for the observed patterns of population genetic diversity and structure of *E. superba* populations in the WAP

region is multiple centers of reproduction and recruitment, utilizing coastal, shelf, and offshore areas. Further use of allele frequencies at SNP sites in mitochondrial genes is warranted for studies of population genetic diversity and structure of high gene flow species at regional to sub-regional spatial scales.

All of the molecular information provided in this work will serve as the basis to answer questions related to *S. thompsoni* and its genetic machinery in relation to different environmental conditions. If in the near future controlled experiments in the laboratory are an option for this species, sequencing of genes known to be important in stress responses and associated biochemical pathways for samples subjected to different conditions would provide answers and predictions on how salps respond to environmental variation over a range of time scales. On another hand, characterization of expression for genes related to reproduction will help understand the mechanism of protogynous hermaphroditism in *S. thompsoni.* Overall, the results of my doctoral research provide a useful database of genes that can aid in understanding the molecular mechanisms *S. thompsoni* uses to thrive under extreme environmental conditions and provide a better understanding of the species' physiology, reproduction, distribution and environment preferences.

References:

Abrusán G, Grundmann N, DeMester L, Makalowski W. (2009). TEclass—a tool for automated classification of unknown eukaryotic transposable elements. Bioinformatics 25: 1329-1330. doi: 10.1093/bioinformatics/btp084.

Alldredge, A.L. and Madin, L.P. (1982). Pelagic Tunicates: Unique Herbivores in the Marine Plankton. *BioScience* 32, 655-663.

Arar, E.J. and G.B. Collins. (1997): EPA Method 445.0: In Vitro Determination of Chlorophyll a and Pheophytin a in Marine and Freshwater Algae by Fluorescence. U.S. EnvironmentalProtection Agency, Washington, DC, EPA-600-R-97-072, 22pp.

Ashjian, C.J., Davis, C.S., Gallager, S.M., Wiebe, P.H., Lawson, G.L., (2008). Distribution of larval krill and zooplankton in association with hydrography in Marguerite Bay, Antarctic Peninsula, in austral fall and winter 2001 described using the Video Plankton Recorder. Deep-Sea Res. II 55: 455-471.

Atkinson, A., V. Siegel, E.A. Pakhomov, P. Rothery, V. Loeb, R.M. Ross, L.B. Quetin, K. Schmidt, P. Fretwell, E.J. Murphy, G.A. Tarling, Fleming A.H., (2008). Oceanic circumpolar habitats of Antarctic krill. Mar. Ecol. Progr. Ser. 362: 1-23.

Atkinson, A., V. Siegel, E. Pakhomov and Rothery, P., (2004). Long-term decline in krill stock and increase in salps within the Southern Ocean. Nature 432(7013): 100-103.

Avise, J.C. (1994) *Molecular Markers, Natural History, and Evolution*. Chapman & Hall, Inc., New York.

Avise, J.C. (2000) *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA. 447 pp.

Ayala, F.J. and Valentine, J.W., (1979). Genetic variability in the pelagic environment: a paradox. Ecology, 60: 24-29.

Atkinson, A., Whitehouse, M.J., Priddle, J., Cripps, G.C., Ward, P., Brandon, M.A., (2001). South Georgia Antarctica: a productive, cold water, pelagic ecosystem. Marine Ecol. Progr. Ser. 216: 279-308.

Baker, A. d. C., M. R. Clarke and M. J. Harris, (1973). "The N.I.O. combination net (RMT 1 + 8) and further developments of rectangular midwater trawls." Journal of the Marine Biological Association of the United Kingdom 53(01): 167-184.

Bandelt, H.J., Forster, P., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16, 37-48.

Baum, S., (2010). Bransfield Strait. Retrieved from http://www.eoearth.org/view/article/150761

Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57, 289-300.

Bluthgen, N., Kielbasa, S.M., Cajavec, B., and Herzel, H. (2004) HOMGL-comparing genelists across species and with different accession numbers. *Bioinformatics* 20, 125-126.

Brinton, E. (1991) Distribution and population structure of immature and adult *Euphausia superba* in the western Bransfield Strait region during the 1986-87 summer. Deep Sea Res. II 38: 1169-1193.

Bucklin A., (2011); http://data.bcodmo.org/LMG11-10/LMG11-10_Cruise_Report_06dec11.pdf

Bucklin, A. and P.H. Wiebe (1986) Genetic heterogeneity in euphausiid populations: *Euphausia krohnii* and *Nematocelis megalops* in North Atlantic Slope Water. Limnol. Oceanogr. 31: 1346-1352.

Bucklin, A. , S.B. Smolenack, A.M. Bentley, and P.H. Wiebe (1997) Gene flow patterns of the euphausiid, *Meganyctiphanes norvegica*, in the N. Atlantic based on DNA sequences for mitochondrial cytochrome oxidase I and cytochrome b. J. Plank. Res. 19:1763-1781.

Bucklin, A. (2000) Methods for Population Genetic Analysis of Zooplankton. Chapter 11 in: *The Zooplankton Methodology Manual*, International Council for the Exploration of the Sea. Academic Press, London. Pp. 533 - 570.

Bucklin, A., P.H. Wiebe, S.B. Smolenack, and M.E. Clarke (2002) Integrated biochemical, molecular genetic, and bioacoustical analysis of mesoscale variability of the euphausiid *Nematoscelis difficilis* in the California Current. Deep-Sea Res. 49: 437-462.

Bucklin, A., P.H. Wiebe, S.B. Smolenack, N.J. Copley, J.G. Beaudet, K.G. Bonner, J. Färber Lorda, and J.J. Pierson (2007) DNA barcodes for species identification of euphausiids (Euphausiacea, Crustacea). J. Plankton Res. 29: 483-493.

Capella, J.E., R.M. Ross, L.B. Quetin, and E.E. Hofmann (1992) A note on the thermal structure of the upper ocean in the Bransfield Strait-South Shetland Islands region. Deep-Sea Res. I 39:1221–1229.

Clarke, A., E.J. Murphy, M.P. Meredith, J.C. King, L.S. Peck, D.K.A. Barnes and R.C. Smith (2007) Climate change and the marine ecosystem of the western Antarctic Peninsula. Phil. Trans. R. Soc. B 362: 149–166.

Clarke, KR, Gorley, RN, (2006). PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.

Conesa, A. and Gotz, S. (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International journal of plant genomics* 2008, 619832.

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., and Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674-3676.

Daponte, M.C., Capitanio, F.L., and Esnal, G.B. (2001) A mechanism for swarming in the tunicate Salpa thompsoni (Foxton, 1961). *Antarctic Science* 13, 240-245.

Deibel, D. and Paffenhofer, G.A. (2009) Predictability of patches of neritic salps and doliolids (Tunicata, Thaliacea). *Journal of Plankton Research* 31, 1571-1579.

Denoeud, F., Henriet, S., Mungpakdee, S., Aury, J.-M., Da Silva, C., Brinkmann, H., Mikhaleva, J., Olsen, L.C., Jubin, C., Cañestro, C., Bouquet, J.-M., Danks, G., Poulain, J., Campsteijn, Konrad, A., Singh, S., Jensen, M.F., Cong, E.H., Eikeseth-Otteraa, H., Noel, B., Anthouard, V., Porcel, B.M., Kachouri-Lafond, R., Nishino, A., Ugolini, M., Chourrout, P., Nishida, H., Aasland, R., Huzurbazar, S., Westhof, E., Delsuc, F., Lehrach, H., Reinhardt, R., Weissenbach, J., Roy, S.W., Artiguenave, F., Postlethwait, J.H., Manak, J.R., Thompson, E.M., Jaillon, O., Du Pasquier, L., Boudinot, P., Liberles, D.A., Volff, J., N., Philippe, H., Lenhard, B., Crollius, H.R., Wincker, P., and Chourrout, D. (2010) Plasticity of Animal Genome Architecture Unmasked by Rapid Evolution of a Pelagic Tunicate. *Science* 330, 1381-1385

Dehal, P., Satou, Y., Campbell, R.K., Chapman, J., Degnan, B., De Tomaso, A., Davidson, B., Di Gregorio, A., Gelpke, M., Goodstein, D.M., Harafuji, N., Hastings, K.E., Ho, I., Hotta, K., Huang, W., Kawashima, T., Lemaire, P., Martinez, D., Meinertzhagen, I.A., Necula, S., Nonaka, M., Putnam, N., Rash, S., Saiga, H., Satake, M., Terry, A., Yamada, L., Wang, H.G., Awazu, S., Azumi, K., Boore, J., Branno, M., Chin-Bow, S., DeSantis, R., Doyle, S., Francino, P., Keys, D.N., Haga, S., Hayashi, H., Hino, K., Imai, K.S., Inaba, K., Kano, S., Kobayashi, K., Kobayashi, M., Lee, B.I., Makabe, K.W., Manohar, C., Matassi, G., Medina, M., Mochizuki, Y., Mount, S., Morishita, T., Miura, S., Nakayama,

A., Nishizaka, S., Nomoto, H., Ohta, F., Oishi, K., Rigoutsos, I., Sano, M., Sasaki, A., Sasakura, Y., Shoguchi, E., Shin-i, T., Spagnuolo, A., Stainier, D., Suzuki, M.M., Tassy, O., Takatori, N., Tokuoka, M., Yagi, K., Yoshizaki, F., Wada, S., Zhang, C., Hyatt, P.D., Larimer, F., Detter, C., Doggett, N., Glavina, T., Hawkins, T., Richardson, P., Lucas, S., Kohara, Y., Levine, M., Satoh, N., and Rokhsar, D.S. (2002) The draft genome of Ciona intestinalis: insights into chordate and vertebrate origins. *Science* 298, 2157-2167.

Dierssen, H. M. (2010). Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate. Proc. Natl. Acad. Sci. 107: 17073–17078.

Dierssen, H. M., R. C. Smith, and M. Vernet. (2000). Optimizing models for remotely estimating primary production in Antarctic coastal waters. Antarctic. Sci. 12: 20–32.

Dierssen, H., and R. C. Smith. (2000). Bio-optical properties and remote sensing ocean color algorithms for Antarctic Peninsula waters. J. Geophys. Res. 105: 26301–26312.

Dubischar, C. D. and U. V. Bathmann (1997). "Grazing impact of copepods and salps on phytoplankton in the Atlantic sector of the Southern ocean." Deep Sea Research Part II: Topical Studies in Oceanography 44(1-2): 415-433.

Ducklow, H. W., K. Baker, D. G. Martinson, L. B. Quetin, R. M. Ross, R. C. Smith, S. E. Stammerjohn, M. Vernet and W. Fraser (2007). "Marine pelagic ecosystems: the West Antarctic Peninsula." Philosophical Transactions of the Royal Society B: Biological Sciences 362(1477): 67-94.

Ducklow, H. W., K. Baker, W.R. Fraser, D.G. Martinson, L.B. Quetin, R.M. Ross, R.C. Smith, S. Stammerjohn and M. Vernet (2007) Marine ecosystems: the West Antarctic Peninsula. Phil. Trans. R. Soc. B 362: 67–94.

Excoffier, L., P.E. Smouse, and J.M. Quattro (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479-491.

Excoffier, L. G. Laval, and S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50.

Fach, B.A., Hofmann, E.E., Murphy, E.J. (2002) Modeling studies of Antarctic krill *Euphausia superba* survival during transport across the Scotia Sea. Marine Ecol. Progr. Ser.231: 187-203.

Fahrbach E. (2011). The Expedition of the Research Vessel "Polarstern" to the Antarctic in 2010/11 (ANT-XXVII/2) http://epic.awi.de/29981/1/Fah2011b.pdf

Fevolden, S.E. and R. Schneppenheim (1989) Genetic homogeneity of krill (*Euphausia superba* Dana) in the Southern Ocean. Polar Biol. 9: 533-539.

Fevolden, S.E. (1988) Biochemical genetics and population structure of *Euphausia superba.* Comp. Biochem. Physiol. B 90B: 507-513.

Fevolden, S.E. (1986) Genetic variation of *Euphausia superba* Dana in the Antarctic Peninsula waters. Sarsia 71: 169-175.

Fevolden, S.E. (1984) Biotic and physical environmental impact on genetic variation of krill. J. Crust. Biol., vol. 4 (suppl.): 206-223

Fevolden, S.E. and F.J. Ayala (1981) Enzyme polymorphism in Antarctic krill (Euphausiacea); genetic variation between populations and species. Sarsia 66:167-173.

Flores, H., Atkinson, A., Kawaguchi, S., Krafft, B.A., Milinevsky, G., Nicol, S., Reiss, C., Tarling, G.A., Werner, R., Bravo Rebolledo, E., Cirelli, V., Cuzin-Roudy, J., Fielding, S., van Franeker, J.A., Groeneveld, J.J., Haraldsson, M., Lombana, A., Marschoff, E., Meyer, B., Pakhomov, E.A., Van de Putte, A.P., Rombolá, E., Schmidt, K., Siegel, V., Teschke, M., Tonkes, H., Toullec, J.Y., Trathan, P.N., Tremblay, N., and Werner, T. (2012) Impact of climate change on Antarctic krill. *Marine Ecology Progress Series* 458, 1-19.

Folmer, O., M. Black, W. Hoen, R. Lutz, and R. Vrijenhoek (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molec. Mar. Biol. Biotech. 3, 294-299.

Fu L, Niu B, Zhu Z, Wu S, Li W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28: 3150-3152. doi: 10.1093/bioinformatics/bts565

Galindo, H.M., D.B. Olson, and S.R. Palumbi (2006) Seascape genetics: a coupled oceanographic-genetic model predicts population structure of Caribbean corals Current Biology 16: 1622–1626.

Gallardo-Escárate, C., Valenzuela-Muñoz, V., and Nuñez-Acuña, G. (2014) RNA-Seq Analysis Using *De Novo* Transcriptome Assembly as a Reference for the Salmon Louse *Caligus rogercresseyi*. *PLoS ONE* 9, e92239.

Gibbs, R.A., P.-N. Nguyen, and C.T. Caskey (1981) Detection of a single DNA base differences by competitive oligonucleotide priming. Nucleic Acids Res. 17: 2437-2448.

Glenn, T.C. (2011) Field guide to next-generation DNA sequencers. *Mol Ecol Resour* 11, 759-769.

Grant, W.S. (1983) Population genetics of krill and comparison with other marine organisms. Ber. Polarforsch./Rep. Polar Res., no. 4: 246)266.

Gurr, J. and V. Siegel (1994) Benthopelagic aggregations of krill (*Euphausia superba*) on the deeper shelf of the Weddell Sea (Antarctic). Deep-Sea Res. 41: 169-178.

Hedgecock, D., P.H. Barber, and S. Edmands (2007) Genetic approaches to measuring connectivity. Oceanography 20: 70-79.

Hill, R.S., Allen, L.D., Bucklin A. (2001) Multiplexed species-specific PCR protocol to discriminate four N. Atlantic *Calanus* species, with an mtCOI gene tree for then *Calanus* species. Marine Biology 139: 279-287.

Hofmann, E.E., P.H. Wiebe, D.P. Costa, and J.J. Torres (2008) Introduction to dynamics of plankton, krill, and predators in relation to environmental features of the western Antarctic Peninsula and related areas: SO GLOBEC Part II. Deep-Sea Res. II: 55: 269-270.

Hofmann, E.E., P.H. Wiebe, D.P. Costa, and J.J. Torres (2004) An Overview of the Southern Ocean Global Ocean Ecosystems Dynamics Program. Deep-Sea Res. II 51: 1921-1924.

Hofmann, E.E., and J.M. Klinck (1998) Thermohaline variability of the waters overlying the west Antarctic Peninsula continental shelf. In: *Ocean, Ice and Atmosphere: Interactions at the Antarctic Continental Margin*, S. Jacobs and R. Weiss (Eds.), pp. 67-81, Antarctic Research Series 75, AGU, Washington, D.C.

Hofmann, E.E., J.M. Klinck, C.M. Lascara and D.A. Smith (1996) Water mass distribution and circulation west of the Antarctic Peninsula and including Bransfield Strait. In: *Foundations for Ecological Research West of the Antarctic Peninsula*, R.M. Ross, E.E. Hofmann, and L.B. Quetin (Eds.), Antarctic Research Series 70, AGU, Washington, D.C. Pp. 61-80.

Hofmann, E.E., Capella, J.E., Ross, R.M., Quetin, L.M. (1992) Models of the early life history of *Euphausia superba*-Part I. Time and temperature dependence during the descent-ascent cycle. Deep-Sea Res. 39: 1177-1200.

Holland, L.Z. and Gibson-Brown, J.J. (2003) The Ciona intestinalis genome: when the constraints are off. *Bioessays* 25, 529-532.

Holt C, Yandell M (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics 12: 491.

Hunt, B. P. V., E. A. Pakhomov and B. G. Trotsenko (2007). "The macrozooplankton of the Cosmonaut Sea, east Antarctica (30°E–60°E), 1987–1990." Deep Sea Research Part I: Oceanographic Research Papers 54(7): 1042-1069.

Huntley,M.E., Brinton, E. (1991) Mesoscale variation in growth and early development of *Euphausia superba* Dana in the western Bransfield Strait region. Deep-Sea Res. II 38, 1213-1240.

Huntley, M., P. Sykes and V. Marin (1989). "Biometry and trophodynamics of Salpa thompsoni foxton (Tunicata: Thaliacea) near the Antarctic Peninsula in austral summer, 1983–1984." Polar Biology 10(1): 59-70.

Issacs, J. E. and L. W. Kidd (1953). "Isaacs-Kidd Midwater Trawl." University of California, Scripps Institute of Oceanography, Final Report 1, 53-3: 18.

Jarman, S.N., N.J. Gales, M. Tierney, P.C. Gill, and N.G. Elliott (2002) A DNA-based method for identification of krill species and its application to analysing the diet of marine vertebrate predators. Moled. Ecol. 11: 2679-2690.

Jenkins A. and Jacobs, S (2006) Circulation and melting beneath George VI Ice Shelf. Antarctica. FRISP Report No. 17 (Forum for Research into Ice Shelf Processes compiled by Lars H. Smedsrud). page 55-66.

Johansson, M.L. (2009) Next Generation Sequencing in Nonmodel Organisms: Has the Future Arrived? *J Hered* 100, 807-.

Jones, D. L. (2014). Fathom Toolbox for Matlab: software for multivariate ecological and oceanographic data analysis. College of Marine Science, University of South Florida, St. Petersburg, FL, USA. Available from: http://www.marine.usf.edu/user/djones/

Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., et al. (2007) The medaka draft genome and insights into vertebrate genome evolution. *Nature* 447, 714-719.

Kasamatsu N., S. Kawaguchi, S. Watanabe, T. Odate and M. Fukuchi (2004) Possible impacts of zooplankton grazing on dimethylsulfide production in the Antarctic Ocean. Can. J. Fish. Aquatic Sci. 61: 736–743.

Kawada, T., Ogasawara, M., Sekiguchi, T., Aoyama, M., Hotta, K., Oka, K., and Satake, H. (2011) Peptidomic analysis of the central nervous system of the protochordate, *Ciona intestinalis*: homologs and prototypes of vertebrate peptides and novel peptides. *Endocrinology* 152, 2416-2427.

Korf I. (2004). Gene finding in novel genomes. BMC Bioinformatics 5: 59.

Kawaguchi, S., V. Siegel, F. Litvinov, V. Loeb and J. Watckins (2004) Salp distribution and size composition in the Atlantic sector of the Southern Ocean. Deep-Sea Res. II 51: 1369-1381.

Klinck, J.M., Hofmann, E.E., Beardsley, R.C., Salihoglu, B., Howard, S. (2004) Water mass properties and circulation on the west Antarctic Peninsula continental shelf in Austral fall and winter 2001. Deep Sea Research II 51: 1925-1946.

Kokubun, N., Kim, J.-H., and Takahashi, A. (2013) Proximity of krill and salps in an Antarctic coastal ecosystem: evidence from penguin-mounted cameras. *Polar Biology* 36, 1857-1864.

Kozlowski, W. A., D. Deutschman, I. Garibotti, C. Trees, and M. Vernet. (2011). An evaluation of the application of CHEMTAX to Antarctic coastal pigment data. Deep Sea Res. Part Oceanogr. Res. Pap. 58: 350–364.

Kuehl, S., and R. Schneppenheim (1986) Electrophoretic investigation of genetic variation in two krill species *Euphausia superba* and *E. crystallorophias* (Euphausiidae). Polar Biol., vol. 6: 17.

Lascara, C.M., E.E. Hofmann, R.R. Ross and L.B. Quetin (1999) Seasonal variability in the distribution of Antarctic krill, *Euphausia superba*, west of the Antarctic Peninsula, Deep Sea Res. 46: 925-949,.

Lawson G.L., P.H. Wiebe, C.J. Ashjian, and T.K. Stanton (2008) Euphausiid distribution along the Western Antarctic Peninsula-Part B: Distribution of euphausiid aggregations and biomass, and associations with environmental factors. Deep-Sea Research II 55: 432-454.

Lee, C.I., Pakhomov, E., Atkinson, A., and Siegel, V. (2010) Long-Term Relationships between the Marine Environment, Krill and Salps in the Southern Ocean. *Journal of Marine Biology* 2010.

Lindblad-Toh, K., E. Winchester, M.J. Daly, D.G. Wang, J.N. Hirschhorn, J.P Laviolette, K. Ardlie, D.E. Reich, E. Robinson, P. Sklar, N. Shah, D. Thomas, J.B. Fan, R. Gingeras, J. Warrington, N. Patil, T.J. Hudson, and E.S. Lander (2000) Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. Nat. Genet. 24: 381-386.

Lizotte, M.P. (2001) The Contributions of Sea Ice Algae to Antarctic Marine Primary Production. American Zoologist 41, 57-73.

Loeb, V. J. and J. A. Santora (2012). "Population dynamics of Salpa thompsoni near the Antarctic Peninsula: Growth rates and interannual variations in reproductive activity (1993–2009)." Progress In Oceanography 96(1): 93-107.

Loeb, V., E. E. Hofmann, J. M. Klinck and O. Holm-Hansen (2010). "Hydrographic control of the marine ecosystem in the South Shetland-Elephant Island and Bransfield Strait region." Deep Sea Research Part II: Topical Studies in Oceanography 57(7–8): 519-542.

Loeb, V.J., E.E. Hofmann, J.M. Klinck, O. Holm-Hansen, and W.B. White (2009) ENSO and variability of the Antarctic Peninsula pelagic marine ecosystem. Antarctic Sci. 21: 135-148.

Loeb, V. J., E. E. Hofmann, J. M. Klinck, O. Holm-Hansen and W. B. White (2008). "ENSO and variability of the Antarctic Peninsula pelagic marine ecosystem." Antarctic Science 21(02): 135.

Loeb, V. (2007). Environmental Variability and the Antarctic Marine Ecosystem. The Impact of Environmental Variability on Ecological Systems. D. A. Vasseur and K. S. McCann, Springer Netherlands. **2:** 197-225

Loeb, V., V. Siegel, O. Holm-Hansen, R. Hewitt, W. Fraser, W. Trivelpiece and S. Trivelpiece (1997). "Effects of sea-ice extent and krill or salp dominance on the Antarctic food web." Nature 387(6636): 897-900.

MacDonald, C.M., and R. Schneppenheim (1983) Breeding structure and stock identity in the Antarctic krill *Euphausia superba* Dana. Ber. Polarforsch./Rep. Polar Res. 4: 240-245.

MacDonald, C.M., R. Williams, and M. Adams (1986) Genetic variation and population structure of krill (*Euphausia superba* Dana) from the Prydz Bay region of Antarctic waters. Polar Biol. 6: 233-236.

Machida, R.J., M.U. Miya, M.M. Yamauchi, M. Nishida, and S. Nishida (2004) Organization of the mitochondrial genome of Antarctic krill *Euphausia superba* (Crustacea: Malacostraca). Marine Biotech. 6: 238-250.

Mackintosh, N. (1973) Distribution of post-larval krill in the Antarctic, Discovery Reports 36: 95–156.

Madin, L.P. and Deibel, D. (1998) Fedind and energetics of Thaliacean. In *The biology of Pelagic Tunicates* (Bone, Q., ed), pp. 81-103, Oxfrod University Press.

Malloy, K.D., M.A. Holman, D. Mitchell, and H.W. Detrich III (1997) Solar UVB-induced DNA damage and photoenzymatic DNA repair in Antarctic zooplankton. Proc. Natl. Acad. Sci. USA 94:1258-1263.

Manel, S. M.K. Schwartz, G. Luikart, and P. Taberlet (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol. Evol. 18: 189-197.

Mare, W. (2009). "Changes in Antarctic sea-ice extent from direct historical observations and whaling records." Climatic Change 92(3-4): 461-493.

Marquez, M.E.I. (1987) Enzyme polymorphism in Antarctic krill (*Euphausia superba*); the phosphoglucomutase locus. Contrib. Cient. Inst. Antart. Argent. No. 335, 16 pp.

Marr, J.W.S. (1962) The natural history and geography of the Antarctic krill (*Euphausia superba* Dana), Discovery Report 32: 33-464.

Massom, R., P. Reid, S. Stammerjohn, B. Raymond, A. Fraser and S. Ushio (2013). "Change and Variability in East Antarctic Sea Ice Seasonality, 1979/80–2009/10." PLoS ONE 8(5): e64756.

Meijers, A. J. S., A. Klocker, N. L. Bindoff, G. D. Williams and S. J. Marsland (2010). "The circulation and water masses of the Antarctic shelf and continental slope between 30 and." Deep Sea Research Part II: Topical Studies in Oceanography 57(9–10): 723-737.

Meyer, E., Aglyamova, G.V., and Matz, M.V. (2011) Profiling gene expression responses of coral larvae (Acropora millepora) to elevated temperature and settlement inducers using a novel RNA-Seq procedure. Molecular Ecology 20, 3599-3616.

Moffat Varas, C. F. (2007). Ocean circulation and dynamics on the west Antarctic Peninsula continental shelf. Ph.D. thesis, Massachusetts Institute of Technology.

Moffat, C., R. C. Beardsley, B. Owens and N. van Lipzig (2008). "A first description of the Antarctic Peninsula Coastal Current." Deep Sea Research Part II: Topical Studies in Oceanography 55(3–4): 277-293.

Moline, M. A., H. Claustre, T. K. Frazer, O. Schofield and M. Vernet (2004). "Alteration of the food web along the Antarctic Peninsula in response to a regional warming trend." Global Change Biology 10(12): 1973-1980.

Montes-Hugo, M., Doney, S.C., Ducklow, H.W., Fraser, W., Martinson, D., Stammerjohn, S.E., and Schofield, O. (2009) Recent Changes in Phytoplankton Communities Associated with Rapid Regional Climate Change Along the Western Antarctic Peninsula. *Science* 323, 1470-1473.

Ngan, P., V. Gomes, H. Suzuki, and M.J.A.C.R. Passos (1990) Preliminary study on chromosomes of Antarctic krill, *Euphausia superba* Dana. Polar Biol. 10: 149-150.

Nicholson G., Smith A.V., Jonsson F., Gustafsson O., Stefansson K. and Donelly P. (2002) Assessing population differentiation and isolation from single-nucleotide polymorphism data. J. R. Statist. Soc. B 64: 695–715.

Nicol, S. and B. Raymond (2012). Pelagic Ecosystems in the Waters off East Antarctica (30° E–150° E). Antarctic Ecosystems, John Wiley & Sons, Ltd: 243-254.

Nishikawa, J., M. Naganobu, T. Ichii, H. Ishii, M. Terazaki and K. Kawaguchi (1995). "Distribution of salps near the South Shetland Islands during austral summer, 1990–1991 with special reference to krill distribution." Polar Biology 15(1): 31-39.

Ohshima, K.I., Takizawa, T., Ushio, S., and Kawamura, T. (1996) Seasonal variations of the Antarctic coastal ocean in the vicinity of Lützow-Holm Bay. Journal of Geophysical Research: Oceans 101, 20617-20628

Ohshima, K. I., T. Kawamura, T. Takizawa, S. Ushio and T. Miyakawa (2000). "Current variability under landfast sea ice in Lützow-Holm Bay, Antarctica." Journal of Geophysical Research: Oceans 105(C7): 17121-17132.

Ono, A., Ishimaru, T., and Tanaka, Y. (2010) Distribution and population structure of salps off Adelie Land in the Southern Ocean during austral summer, 2003 and 2005. *La mer* 48, 55-70.

Orsi, A. H., T. Whitworth Iii and W. D. Nowlin Jr (1995). "On the meridional extent and fronts of the Antarctic Circumpolar Current." Deep Sea Research Part I: Oceanographic Research Papers 42(5): 641-673.

Oshima, T., K. Watanabe, K. Asai, and Y. Kuchino (1981) Chemical structure and thermal properties of initiator tRNA from Euphausia superba in comparison with those of other eukaryotic initiator of tRNAs. J. Biochem., Tokyo 90: 1259-1266.

Pakhomov, E.A., Fuentes, V., Schloss, I., Atencio, A., and Esnal, G.B. (2003) Beaching of the tunicate Salpa thompsoni at high levels of suspended particulate matter in the Southern Ocean. *Polar Biology* 26, 427-431.

Pakhomov, E. A. (2004). "Salp/krill interactions in the eastern Atlantic sector of the Southern Ocean." Deep Sea Research Part II: Topical Studies in Oceanography 51(22-24): 2645-2660.

Pakhomov, E. A. and P. W. Froneman (2004). "Zooplankton dynamics in the eastern Atlantic sector of the Southern Ocean during the austral summer 1997/1998—Part 2: Grazing impact." Deep Sea Research Part II: Topical Studies in Oceanography 51(22-24): 2617-2631.

Pakhomov, E., Dubischar, C., Strass, V., Brichta, M., and Bathmann, U. (2006) The tunicate Salpa thompsoni ecology in the Southern Ocean. I. Distribution, biomass, demography and feeding ecophysiology. *Marine Biology* 149, 609-623.

Pakhomov, E.A., Dubischar, C.D., Hunt, B.P.V., Strass, V., Cisewski, B., Siegel, V., Bathmann, U. (2011) Biology and life cycles of pelagic tunicates in the Lazarev Sea, Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* 58, 1677-1689.

Papetti, C., L. Zane, E. Bortolotto, A. Bucklin, and T. Patarnello (2005) Genetic differentiation and local temporal stability of population structure in the euphausiid *Meganyctiphanes norvegica*. Mar. Ecol. Prog. Ser. 289: 225 – 235.

Park, C. and J. H. Wormuth (1993). "Distribution of antarctic zooplankton around Elephant Island during the austral summers of 1988, 1989, and 1990." Polar Biology 13(4): 215-225.

Park, Y.-H., E. Charriaud, P. Craneguy and A. Kartavtseff (2001). "Fronts, transport, and Weddell Gyre at 30°E between Africa and Antarctica." Journal of Geophysical Research: Oceans 106(C2): 2857-2879.

Patarnello T., L. Bargelloni, V. Varotto, and B. Battaglia (1996) Krill evolution and the Antarctic ocean currents: evidence of vicariant speciation as inferred by molecular data. Mar. Biol. 126: 603-608.

Perissinotto, R. and E. A. Pakhomov (1998). "The trophic role of the tunicate Salpa thompsoni in the Antarctic marine ecosystem." Journal of Marine Systems 17(1-4): 361-374.

Phillips, B., Kremer, P., and Madin, L. (2009) Defecation by *Salpa thompsoni*; and its contribution to vertical flux in the Southern Ocean. Marine Biology 156, 455-467.

Quetin, L.B. and R.M. Ross (2001) Environmental variability and its impact on the reproductive cycle of Antarctic krill. American Zoologist: 41:74–89.

Quetin, L. B., R. M. Ross, T. K. Frazer and K. L. Haberman (1996). Factors affecting distribution and abundance of zooplankton, with an emphasis on Antarctic krill, *Euphausia superba*. Foundations for Ecological Research West of the Antarctic Peninsula. Washington, DC, AGU. 70: 357-371.

Quinn, N.L., Levenkova, N., Chow, W., Bouffard, P., Boroevich, K.A., Knight, J.R., Jarvie, T.P., Lubieniecki, K.P., Desany, B.A., Koop, B.F., Harkins, T.T., and Davidson, W.S. (2008) Assessing the feasibility of GS FLX Pyrosequencing for sequencing the Atlantic salmon genome. *BMC Genomics* 9, 404.

Quintans, B.A., V. lvarez-Iglesias, A.C. Salas, M.V. Phillips, and L.A. Carracedo (2004) Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing, Forensic Sci. Int. 140, 251–257.

Reid, P., S. Stammerjohn, R. Massom, T. Scambos, and J. Lieser. (2015, in press). The record 2013 Southern Hemisphere sea-ice extent maximum. Annals of Glaciology 56 (69), doi:10.3189/2015AoG69A892.

Rintoul, S. R. (2009). Antarctic Circumpolar Current. Encyclopedia of Ocean Sciences (Second Edition). H. S. Editors-in-Chief: John, K. T. Karl and A. T. Steve. Oxford, Academic Press: 178-190.

Roe, H. S. J., & Shale, D. M. (1979). A new multiple rectangular midwater trawl (RMT 1+8M) and some modifications to the Institute of Oceanographic Sciences' RMT 1+8. Marine Biology, 50, 283–288.

Rohardt, G. (2011). Continuous thermosalinograph oceanography along POLARSTERN cruise track ANT-XXVII/2. H. C. f. P. a. M. R. Alfred Wegener Institute, Bremerhaven.

Ross, R. M., L. B. Quetin, and C. M. Lascara (1996) Distribution of Antarctic krill and dominant zooplankton west of the Antarctic Peninsula. In R. M. Ross, E. E. Hofmann, and L. B. Quetin (Eds.), *Foundations for ecological research west of the Antarctic Peninsula*. Antarctic Research Series, Vol. 70, pp. 199–217. American Geophysical Union, Washington, D.C.

Ross, R.M., L.B. Quetin, K.S. Baker, M. Vernet, and R.C. Smith (2000) Growth limitation in young *Euphausia superba* under field conditions. Limnol. Oceanogr. 45: 31-43.

Rozas, J., J. C. Sanchez-DelBarrio, et al. (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19(18): 2496-2497.

Saba, G.K., Fraser, W.R., Saba, V.S., Iannuzzi, R.A., Coleman, K.E., Doney, S.C., Ducklow, H.W., Martinson, D.G., Miles, T.N., Patterson-Fraser, D.L., Stammerjohn, S.E., Steinberg, D.K.,Schofield, O.M. (2014) Winter and spring controls on the summer food web of the coastal West Antarctic Peninsula. *Nat Commun* 5.

Salmela L, Schröder J. (2011). Correcting errors in short reads by multiple alignments. Bioinformatics 27: 1455-1461.

Sangrà, P., C. Gordo, M. Hernández-Arencibia, A. Marrero-Díaz, A. Rodríguez-Santana, A. Stegner, A. Martínez-Marrero, J. L. Pelegrí and T. Pichon (2011). "The Bransfield current system." Deep Sea Research Part I: Oceanographic Research Papers 58(4): 390-402.

Schneppenheim, R. and M. MacDonald (1983) Population genetics of krill (*Euphausia superba*). Polarforsch./Rep. Polar Res. 4: 239-240.

Seo, H.-C., Kube, M., Edvardsen, R.B., Jensen, M.F., Beck, A., Spriet, E., Gorsky, G., Thompson, E.M., Lecrach, H., Reinhardt, R., Chourrout, D. (2001) Miniature Genome in the Marine Chordate Oikopleura dioica. *Science* 294, 2506-.

Siegel, V., C. S. Reiss, K. S. Dietrich, M. Haraldsson and G. Rohardt (2013). "Distribution and abundance of Antarctic krill (Euphausia superba) along the Antarctic Peninsula." Deep Sea Research Part I: Oceanographic Research Papers 77(0): 63-74.

Siegel, V. (2005) Distribution and population dynamics of *Euphasia superba:* summary of recent findings. Polar Biol. 29:1-22.

Siegel, V. (1988) A concept of seasonal variation of krill (*Euphausia superba*) distribution and abundance west of the Antarctic Peninsula. In: Sarhage, D.(Ed.), *Antarctic Ocean and Resources Variability*. Springer, Berlin, Heidelberg, pp. 219-230.

Siegel, V. (1995). "Recruitment of Antarctic krill (Euphausia superba) and possible causes for its variability." Mar Ecol Prog Ser 123: 45-56.

Smit AFA, Hubley, R., and Green, P. 1996-2010. RepeatMasker Open-3.0. (http://www.repeatmasker.org/)

Smith, D. A., E. E. Hofmann, J. M. Klinck and C. M. Lascara (1999). "Hydrography and circulation of the West Antarctic Peninsula Continental Shelf." Deep Sea Research Part I: Oceanographic Research Papers 46(6): 925-949.

Smith, S., Bernatchez, L., and Beheregaray, L.B. (2013) RNA-seq analysis reveals extensive transcriptional plasticity to temperature stress in a freshwater fish species. *BMC Genomics* 14, 375.

Smith, S.L. and S.B. Schnack-Schiel (1990) Polar zooplankton. In: W.O. Smith (Ed.) *Polar Oceanography Part B Chemistry, Biology, and Geology*, Academic Press, New York. Pp. 527–598.

Sparrow, M. D., K. J. Heywood, J. Brown and D. P. Stevens (1996). "Current structure of the south Indian Ocean." Journal of Geophysical Research: Oceans 101(C3): 6377-6391.

Spiridinov, V.A., E.-M. Nöthig, M. Schröder and A. Wisotzki (1996) The onset of biological winter in the eastern Weddell Gyre (Antarctica) planktonic community. J. Mar. Systems 9: 211-230.

Stach, T. (2007) Ontogeny of the appendicularian <i>Oikopleura dioica</i> (Tunicata, Chordata) reveals characters similar to ascidian larvae with sessile adults. *Zoomorphology* 126, 203-214.

Stanke M, Morgenstern B. (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Research 33: W465-W467. doi: 10.1093/nar/gki458

Stein, M. (1992) Variability of local upwelling off the Antarctic Peninsula, 1986-1990. Arch. FischWiss. 41: 131-158.

Stern, M., and R.B. Heywood (1994) Antarctic environment - physical oceanography: the Antarctic Peninsula and Southwest Atlantic region of the Southern Ocean. In Southern

Ocean Ecololgy: The BIOMASS Perspective, S. Z. El-Sayed [Ed], Cambridge University Press, New York. Pages 11-24.

Takatori, N., Hotta, K., Mochizuki, Y., Satoh, G., Mitani, Y., Satoh, N., Satou, Y., Takahashi, H. (2004) T-box genes in the ascidian Ciona intestinalis: Characterization of cDNAs and spatial expression. *Developmental Dynamics* 230, 743-753.

Thompson, A. F., K. J. Heywood, S. E. Thorpe, A. H. H. Renner and A. Trasviña (2009). "Surface Circulation at the Tip of the Antarctic Peninsula from Drifters." Journal of Physical Oceanography 39(1): 3-26.

Thorpe, S.E., E.J. Murphy, and J.L. Watkins (2007) Circumpolar connections between Antarctic krill (*Euphausia superba* Dana) populations: Investigating the roles of ocean and sea ice transport. Deep-Sea Res. Part I. 54: 792-810.

USEPA (U.S. Environmental Protection Agency), "ESS Method 310.1: Ortho-Phosphorus, Dissolved Automated, Ascorbic Acid," Environmental Sciences Section Inorganic Chemistry Unit, Wisconsin State Lab of Hygiene, Madison, 1992.

Van, T.V., Kimoto, K., and Murakami, K. (1982) Purification and some properties of two ribonucleases from Euphausia superba. Agric. Biol. Chem. 46: 691-696.

Vera, J.C., Christopher W, W., Howard W, F., Mikko J, F., Douglas L, C., Ilkka, H., and James H, M. (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Molecular Ecology* 17, 1636-1647.

Villarino, G.H., Bombarely, A., Giovannoni, J.J., Scanlon, M.J., and Mattson, N.S. (2014) Transcriptomic Analysis of *Petunia hybrida* in Response to Salt Stress Using High Throughput RNA Sequencing. *PLoS ONE* 9, e94651.

Voskoboynik, A., Neff, N.F., Sahoo, D., Newman, A.M., Pushkarev, D., Koh, W., Passarelli, B., Fan, H. C., Mantalas, G. L., Palmeri, K. J., Ishizuka, K. J., Gissi, C., Griggio, F., Ben-Shlomo, R., Corey, D. M., Penland, L., White, R. A., Weissman, I. L., Quake, S.R. (2013) *The genome sequence of the colonial chordate, Botryllus schlosseri*.

Wang D.G., J.B. Fan, C. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Weir, B.S. and W.G. Hill (2002) Estimating F-statistics. Ann. Rev. Genetics 36:721-750.

Wiebe, P.H., C. Ashjian, S. Gallager, C. Davis, G. Lawson, and N.J. Copley (2004) Using a high powered strobe light to increase the catch of Antarctic krill. Mar. Biol. 144: 493 - 502.

Wiebe, P.H., Morton, A.W., Bradley, A.M., Backus, R.H., Craddock, J.E., Cowles, T.J., Barber, V.A. and Flierl, G.R. (1985) New developments in the MOCNESS, an apparatus for sampling zooplankton and micronekton. Mar. Biol. 87: 313-323.

Wiebe, P.H., Ashjian, C.J., Lawson, G.L., Piñones, A., and Copley, N.J. (2011) Horizontal and vertical distribution of euphausiid species on the Western Antarctic Peninsula U.S. GLOBEC Southern Ocean study site. *Deep Sea Research Part II: Topical Studies in Oceanography* 58, 1630-1651.

Wiebe, P. H., A. W. Morton, A. M. Bradley, R. H. Backus, J. E. Craddock, V. Barber, T. J. Cowles and G. R. Flierl (1985). "New development in the MOCNESS, an apparatus for sampling zooplankton and micronekton." Marine Biology 87(3): 313-323.

Wiebe, P.H., Madin, L.P., Haury, L.R., Harbison, G.R., and Philbin, L.M. (1979) Diel vertical migration bySalpa aspera and its potential for large-scale particulate organic matter transport to the deep-sea. *Marine Biology* 53, 249-255.

Wilhelm J, Pingoud A, Hahn M. (2003). Real-time PCR-based method for the estimation of genome sizes. Nucleic Acids Research 31: e56-e56.

Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M.S. Morris, H. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T.J. Hudson, R. Lipshutz, M. Chee, and E.S. Lander (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280:1077-1082.

Xue, W., Li, J.T., Zhu, Y.P., Hou, G.Y., Kong, X.F., Kuang, Y.Y., and Sun, X.W. (2013) L_RNA_scaffolder: scaffolding genomes with transcripts. *BMC Genomics* 14, 604.

Yabu, H. and A. Kawamura (1984) Chromosomes of *Euphausia superba* Dana. Bull. Plankton Soc. Japan 31:61-63.

Yadetie, F., Butcher, S., Forde, H.E., Campsteijn, C., Bouquet, J.M., Karlsen, O.A., Denoeud, F., Metpally, R., Thompson, E.M., Manak, J.R., Goksoyr, A., and Chourrout, D. (2012) Conservation and divergence of chemical defense system in the tunicate *Oikopleura dioica* revealed by genome wide response to two xenobiotics. BMC Genomics 13, 55.

Yin, J., J. T. Overpeck, S. M. Griffies, A. Hu, J. L. Russell and R. J. Stouffer (2011). "Different magnitudes of projected subsurface ocean warming around Greenland and Antarctica." Nature Geoscience 4(8): 524-528.

Zane L., L. Ostellari, L. Maccatrozzo, L. Bargelloni, B. Battaglia, and T. Patarnello (1998) Molecular evidence for genetic subdivision of Antarctic krill. Proc. R. Soc. Lond. B 265: 2387-2391.

Zane, L., L. Ostellari, L. Maccatrozzo, L. Bargelloni, J. Cuzin-Roudy, F. Buchholz and T. Patarnello (2000) Genetic differentiation in a pelagic crustacean (*Meganyctiphanes norvegica*, Euphausiacea) from the North East Atlantic and the Mediterranean Sea. Mar. Biol. 49: 191-199.

Zane, L, and T. Patarnello (2000) Krill: a model for investigating the effects of ocean currents on the genetic structure of a pelagic organism. Can J. Fish. Aquat. Sci. 57:16-23.

Zhang, J.-Z., and G.A. Berberian, (1997): EPA Method 366.0: Determination of dissolved silicate in estuarine and coastal waters by gas segmented flow colorimetric analysis. U.S. Environmental Protection Agency, Washington, DC, EPA-600-R-97-072, 13 pp.

Zhang, J.-Z., P.B. Ortner, and C.J. Fischer, (1997): EPA Method 353.4: Determination of nitrate and nitrite in estuarine and coastal waters by gas segmented continuous flow colorimetric analysis. U.S. Environmental Protection Agency, Washington, DC, EPA-600-R-97-072, 2pp.

Zimmerman, C.F., C.W. Keefe and J. Bashe. (1997). Determination of Carbon and Nitrogen in Sediments and Particulates of Estuarine/coastal Waters Using Elemental Analysis - Method 440.0. U.S. Environmental Protection Agency, Cincinnati, OH.